

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 19, 2000, 03:36:13 : Search time 2066.03 Seconds
(without alignments)
1024.800 Million cell updates/sec

Title: US-08-978-456-1

Perfect score: 1029
Sequence: 1 ATGGATTATGGATTCAACT.....TAAGTTACGAAGAAGACTGA 1029

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10041192 seqs, 1028798125 residues

Total number of hits satisfying chosen parameters: 20082384

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Pending_Patents_NA.*
1: /cgn2_6/ptodata/2/pna/US082B_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/US06_COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US07_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US081A_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US081B_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US081C_COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US081D_COMB.seq.*
8: /cgn2_6/ptodata/2/pna/US082A_COMB.seq.*
9: /cgn2_6/ptodata/2/pna/US082B_COMB.seq.*
10: /cgn2_6/ptodata/2/pna/US082C_COMB.seq.*
11: /cgn2_6/ptodata/2/pna/US083A_COMB.seq.*
12: /cgn2_6/ptodata/2/pna/US083B_COMB.seq.*
13: /cgn2_6/ptodata/2/pna/US084A_COMB.seq.*
14: /cgn2_6/ptodata/2/pna/US084B_COMB.seq.*
15: /cgn2_6/ptodata/2/pna/US084C_COMB.seq.*
16: /cgn2_6/ptodata/2/pna/US084D_COMB.seq.*
17: /cgn2_6/ptodata/2/pna/US085_COMB.seq.*
18: /cgn2_6/ptodata/2/pna/US085A_COMB.seq.*
19: /cgn2_6/ptodata/2/pna/US085B_COMB.seq.*
20: /cgn2_6/ptodata/2/pna/US086_COMB.seq.*
21: /cgn2_6/ptodata/2/pna/US086A_COMB.seq.*
22: /cgn2_6/ptodata/2/pna/US086B_COMB.seq.*
23: /cgn2_6/ptodata/2/pna/US086C_COMB.seq.*
24: /cgn2_6/ptodata/2/pna/US086D_COMB.seq.*
25: /cgn2_6/ptodata/2/pna/US087A_COMB.seq.*
26: /cgn2_6/ptodata/2/pna/US087B_COMB.seq.*
27: /cgn2_6/ptodata/2/pna/US087C_COMB.seq.*
28: /cgn2_6/ptodata/2/pna/US087D_COMB.seq.*
29: /cgn2_6/ptodata/2/pna/US088A_COMB.seq.*
30: /cgn2_6/ptodata/2/pna/US088B_COMB.seq.*
31: /cgn2_6/ptodata/2/pna/US088C_COMB.seq.*
32: /cgn2_6/ptodata/2/pna/US088D_COMB.seq.*
33: /cgn2_6/ptodata/2/pna/US088E_COMB.seq.*
34: /cgn2_6/ptodata/2/pna/US088F_COMB.seq.*
35: /cgn2_6/ptodata/2/pna/US088G_COMB.seq.*
36: /cgn2_6/ptodata/2/pna/US088H_COMB.seq.*
37: /cgn2_6/ptodata/2/pna/US088I_COMB.seq.*
38: /cgn2_6/ptodata/2/pna/US088J_COMB.seq.*
39: /cgn2_6/ptodata/2/pna/US088K_COMB.seq.*
40: /cgn2_6/ptodata/2/pna/US088L_COMB.seq.*
41: /cgn2_6/ptodata/2/pna/US088M_COMB.seq.*
42: /cgn2_6/ptodata/2/pna/US088N_COMB.seq.*
43: /cgn2_6/ptodata/2/pna/US088O_COMB.seq.*

44: /cgn2_6/ptodata/2/pna/US092B_COMB.seq.*
45: /cgn2_6/ptodata/2/pna/US092C_COMB.seq.*
46: /cgn2_6/ptodata/2/pna/US092D_COMB.seq.*
47: /cgn2_6/ptodata/2/pna/US092E_COMB.seq.*
48: /cgn2_6/ptodata/2/pna/US092F_COMB.seq.*
49: /cgn2_6/ptodata/2/pna/US092G_COMB.seq.*
50: /cgn2_6/ptodata/2/pna/US092H_COMB.seq.*
51: /cgn2_6/ptodata/2/pna/US092I_COMB.seq.*
52: /cgn2_6/ptodata/2/pna/US092J_COMB.seq.*
53: /cgn2_6/ptodata/2/pna/US092K_COMB.seq.*
54: /cgn2_6/ptodata/2/pna/US092L_COMB.seq.*
55: /cgn2_6/ptodata/2/pna/US092M_COMB.seq.*
56: /cgn2_6/ptodata/2/pna/US092N_COMB.seq.*
57: /cgn2_6/ptodata/2/pna/US092O_COMB.seq.*
58: /cgn2_6/ptodata/2/pna/US092P_COMB.seq.*
59: /cgn2_6/ptodata/2/pna/US092Q_COMB.seq.*
60: /cgn2_6/ptodata/2/pna/US092R_COMB.seq.*
61: /cgn2_6/ptodata/2/pna/US092S_COMB.seq.*
62: /cgn2_6/ptodata/2/pna/US092T_COMB.seq.*
63: /cgn2_6/ptodata/2/pna/US092U_COMB.seq.*
64: /cgn2_6/ptodata/2/pna/US092V_COMB.seq.*
65: /cgn2_6/ptodata/2/pna/US092W_COMB.seq.*
66: /cgn2_6/ptodata/2/pna/US092X_COMB.seq.*
67: /cgn2_6/ptodata/2/pna/US092Y_COMB.seq.*
68: /cgn2_6/ptodata/2/pna/US092Z_COMB.seq.*
69: /cgn2_6/ptodata/2/pna/US093A_COMB.seq.*
70: /cgn2_6/ptodata/2/pna/US093B_COMB.seq.*
71: /cgn2_6/ptodata/2/pna/US093C_COMB.seq.*
72: /cgn2_6/ptodata/2/pna/US093D_COMB.seq.*
73: /cgn2_6/ptodata/2/pna/US093E_COMB.seq.*
74: /cgn2_6/ptodata/2/pna/US093F_COMB.seq.*
75: /cgn2_6/ptodata/2/pna/US093G_COMB.seq.*
76: /cgn2_6/ptodata/2/pna/US093H_COMB.seq.*
77: /cgn2_6/ptodata/2/pna/US093I_COMB.seq.*
78: /cgn2_6/ptodata/2/pna/US093J_COMB.seq.*
79: /cgn2_6/ptodata/2/pna/US093K_COMB.seq.*
80: /cgn2_6/ptodata/2/pna/US093L_COMB.seq.*
81: /cgn2_6/ptodata/2/pna/US093M_COMB.seq.*
82: /cgn2_6/ptodata/2/pna/US093N_COMB.seq.*
83: /cgn2_6/ptodata/2/pna/US093O_COMB.seq.*
84: /cgn2_6/ptodata/2/pna/US093P_COMB.seq.*
85: /cgn2_6/ptodata/2/pna/US093Q_COMB.seq.*
86: /cgn2_6/ptodata/2/pna/US093R_COMB.seq.*
87: /cgn2_6/ptodata/2/pna/US093S_COMB.seq.*
88: /cgn2_6/ptodata/2/pna/US093T_COMB.seq.*
89: /cgn2_6/ptodata/2/pna/US093U_COMB.seq.*
90: /cgn2_6/ptodata/2/pna/US093V_COMB.seq.*
91: /cgn2_6/ptodata/2/pna/US093W_COMB.seq.*
92: /cgn2_6/ptodata/2/pna/US093X_COMB.seq.*
93: /cgn2_6/ptodata/2/pna/US093Y_COMB.seq.*
94: /cgn2_6/ptodata/2/pna/US093Z_COMB.seq.*
95: /cgn2_6/ptodata/2/pna/US094A_COMB.seq.*
96: /cgn2_6/ptodata/2/pna/US094B_COMB.seq.*
97: /cgn2_6/ptodata/2/pna/US094C_COMB.seq.*
98: /cgn2_6/ptodata/2/pna/US094D_COMB.seq.*
99: /cgn2_6/ptodata/2/pna/US094E_COMB.seq.*
100: /cgn2_6/ptodata/2/pna/US094F_COMB.seq.*
101: /cgn2_6/ptodata/2/pna/US094G_COMB.seq.*
102: /cgn2_6/ptodata/2/pna/US094H_COMB.seq.*
103: /cgn2_6/ptodata/2/pna/US094I_COMB.seq.*
104: /cgn2_6/ptodata/2/pna/US094J_COMB.seq.*
105: /cgn2_6/ptodata/2/pna/US094K_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1029	100.0	1029	1	PCR-US98-24857-1

2 1029 100.0 1029 103 US-09-376-633-1 Sequence 1, Appl1
3 1029 100.0 3336 1 PCT-US98-24857-7 Sequence 7, Appl1
4 1029 100.0 3336 1 PCT-US98-25068-7 Sequence 7, Appl1
5 1029 100.0 3336 1 PCT-US98-25070-7 Sequence 7, Appl1
6 1029 100.0 3336 1 PCT-US98-25096-5 Sequence 5, Appl1
7 1029 100.0 3336 32 US-08-977-553-5 Sequence 7, Appl1
8 1029 100.0 3336 32 US-08-977-866-7 Sequence 7, Appl1
9 1029 100.0 3336 43 US-09-225-967-7 Sequence 7, Appl1
10 1029 100.0 3336 54 US-09-227-806-7 Sequence 7, Appl1
11 1029 100.0 3336 43 US-09-492-479-7 Sequence 7, Appl1
12 1029 100.0 3336 103 US-09-376-633-7 Sequence 7, Appl1
13 1006.4 97.8 1119 23 US-08-827-356-1376 Sequence 1376, Ap
14 1006.4 97.8 74105 57 US-09-611-529-2079 Sequence 2079, Ap
15 1006.4 97.8 74105 57 US-09-611-529-7446 Sequence 7446, Ap
16 1006.4 97.8 85137 23 US-08-831-156A-107 Sequence 107, Ap
17 1000 97.2 7588 21 US-08-781-986A-142 Sequence 142, Ap
18 1000 97.2 7588 30 US-08-956-171-142 Sequence 142, Ap
19 1000 97.2 7588 30 US-08-956-171B-142 Sequence 142, Ap
20 985.6 95.8 7003 65 US-60-038-697-616 Sequence 616, Ap
21 985.6 95.8 7003 65 US-60-046-714-618 Sequence 618, Ap
22 515.2 50.1 1053 53 US-09-450-969-2343 Sequence 2343, Ap
23 515.2 50.1 8374 70 US-60-068-228-315 Sequence 315, Ap
24 515.2 50.1 8413 65 US-60-038-081-306 Sequence 306, Ap
25 515.2 50.1 8414 67 US-60-046-933-313 Sequence 313, Ap
26 494.6 48.1 509 57 US-09-620-608-372 Sequence 372, Ap
27 494.6 48.1 509 86 US-60-144-883-372 Sequence 372, Ap
28 352.2 34.2 1269 1 PCT-US97-02318-946 Sequence 946, Ap
29 352.2 34.2 1269 1 PCT-US98-24857-3 Sequence 946, Ap
30 352.2 34.2 1269 26 US-08-903-470-946 Sequence 946, Ap
31 352.2 34.2 1269 103 US-09-376-633-3 Sequence 3, Appl1
32 154.4 15.0 31495 57 US-09-663-779-1376 Sequence 1376, Ap
33 151.8 14.8 2507 57 US-09-663-779-829 Sequence 829, Ap
34 143 13.9 258 53 US-09-450-969-2251 Sequence 2251, Ap
35 120.4 11.7 5567 25 US-08-899-241-1 Sequence 1, Appl1
36 115.4 11.2 1101 48 US-09-328-352-1336 Sequence 1336, Ap
37 114.2 11.1 1101 56 US-09-383-110-2144 Sequence 2144, Ap
38 114.2 11.1 14273 31 US-08-961-527-40 Sequence 40, Appl
39 113.8 11.1 15056 63 US-60-029-960-31 Sequence 31, Appl
40 112.6 10.9 1101 1 PCT-US98-25010-1 Sequence 1, Appl1
41 112.6 10.9 1101 32 US-08-979-616-1 Sequence 1, Appl1
42 112.6 10.9 3417 1 PCT-US98-24853-7 Sequence 7, Appl1
43 112.6 10.9 3417 1 PCT-US98-25010-7 Sequence 7, Appl1
44 112.6 10.9 3417 1 PCT-US98-25073-7 Sequence 7, Appl1
45 112.6 10.9 3417 32 US-08-977-555-7 Sequence 7, Appl1

ALIGNMENTS

RESULT 1
PCT-US98-24857-1
Sequence 1, Application PC/PIUS9824857A
GENERAL INFORMATION:
APPLICANT: SmithKline Beecham Corporation
TITLE OF INVENTION: Novel ribg
FILE REFERENCE: P5044-9
CURRENT APPLICATION NUMBER: PCT/US98/24857A
CURRENT FILING DATE: 1998-11-23
EARLIER APPLICATION NUMBER: 08/978,456
EARLIER FILING DATE: 1997-11-25
NUMBER OF SEQ ID NOS: 7
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 1029
TYPE: DNA
ORGANISM: Staphylococcus aureus
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1026)
PCT-US98-24857-1

Query Match

100.0%; Score 1029; DB 1; Length 1029;

Best Local Similarity 100.0%; Pred. No. 4,3e-224;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGATATATGCGATTCACCTTCGCAATATGATGCAACGTCACAGCGTGTATCCACC 60
Db 1 atgattatagattacacttcgcaaatatgataatgataaagtgtaaacaggtgttaaccacc 60
QY 61 GTTGCCCTGTGAGTAAATGAAGAGAGATGTTGGATTGGTGACACTTGAGAAA 120
Db 61 gttagccctgtgagtaaatgaagagatgttggattgggtgacacttgagaaa 120
QY 61 gttagccctgtgagtaaatgaagagatgttggattgggtgacacttgagaaa 120
Db 61 gttagccctgtgagtaaatgaagagatgttggattgggtgacacttgagaaa 120
QY 121 GGTGACAGCATGCGGAGGTTCAACGACTTGATATGCGACAAACAAATGCTGAAGTCCG 180
Db 121 ggtgacagcatgcgaggttcaacgacttgaatgacgcaacaaatgctgaagtgccg 180
QY 181 ACGATTTATATAGCTAGTAGCCCATGAGCATTTGGTTCACACACCACTCGTGTAAAC 240
Db 181 acgatttatatagctagtagcccatgagcatgttggttcacaacacacctgtgtaac 240
QY 241 AAATTTATGTTGATGATAGTAGCAAAAGTAGATACGCAACAAAGCAATTTGCTTAGAC 300
Db 241 aaatttatgttgaatgataagtagatagcaaaagtagatagcaaaagcaatttcgtagac 300
QY 301 ACACATGCTGATGAGAGCGTTACGGGCTCAGGTAATGAGGTTGATGATGATGAA 360
Db 301 acacatgctgatgagagcgttacgggctcaggtatgaggttgaatgaggtatgagtaa 360
QY 361 CGGCATGACATATATCCCAAGACTTTTAAAGCAACAAAGCAACGACGCAATTT 420
Db 361 cggcatgacatattatcccaagactttttaaagcaaacaaagcaacgacgcaaat 420
QY 421 ACACTGAAGATATCGCAAGTTTATGATGATGAACACGCAATGATGACAAAGTCAA 480
Db 421 acactgaagatattcgcaagtttatgattgataacacgcaatgattgacaaagtgcaa 480
QY 481 TGGATTCTAACAAGAGGTTAAACAAGATGCTATAGTTAGACATCGACACGACGA 540
Db 481 tggattcttaacaagaggtttaaacaagatgctatagttagacatcgacacgacga 540
QY 541 GTGTATCTGGAAGAGCTACAGTGAATGATGATGATGATGATGATGATGATGATGATG 600
Db 541 gtgtatctggaagagctacagtggaatgattgattgattgattgattgattgattgatt 600
QY 601 GATGAAAAAACCCTATAAAGATATATGCTGAAGTCTGGAATATTCATTTTATACG 660
Db 601 gatgaaaaaaccttataaagatattatgctgaagtcctggaaattatcatltaacag 660
QY 661 CAATTTATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
Db 661 caatttatcagatgattgattgattgattgattgattgattgattgattgattgattgatt 720
QY 721 AGCATGAACACATATGGAATATTTACTGTAAGTCTGTAAGTCTGTAAGTCTGTAAGTCT 780
Db 721 agcatgaacacatattggaattatcttactgtaagtcctgtaagtcctgtaagtcctgta 780
QY 781 CACATTTATATAAAGAGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 840
Db 781 cacatttatataaagagagtggaagtgaagtgaagtgaagtgaagtgaagtgaagtgaagtga 840
QY 841 TCAGATTTTCATCTAT 900
Db 841 tcagattttcatctatataatataatataatataatataatataatataatataatataat 900
QY 901 GCGGATCTGGAATTTATCAATTTTATCAACAAAGATGATGATGATGATGATGATGATG 960
Db 901 gcggatctggaattttatcaattttatcaacaaagatgattgattgattgattgattgattg 960
QY 961 AACCAATTTGAATTTGTCATTCGAGTATTTAAATGAATGAATGAATGAATGAATGAATGA 1020
Db 961 aaccaatttgaatttgcatttcgagttatataatgaatgaatgaatgaatgaatgaatgaat 1020
QY 1021 AAGAAGTGA 1029
Db 1021 aagaagtga 1029

Db 1021 aaagaagtga 1029

RESULT 2
US-09-376-633-1

GENERAL INFORMATION:
APPLICANT: Palmer, Leslie M.
TITLE OF INVENTION: Novel rldg
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:

COMPUTER READABLE FORM:

```

; SOFTWARE: FastSEQ for Windows Version 2.0
; CHILDREN ADDITION DATA

```

; INFORMATION FOR SEQ ID NO: 1

Query Match	100.0%	Score 1029	DB 103	Length 1029
Best Local Similarity	100.0%	Pred. No. 4.3e-224		
Matches 1029	0	Mismatches 0	Indels 0	Gaps 0

Oy	1	ATGGATTATGGCATTCACCTGCAAAATGGTATCAAGTCAAAACAGGTGTTAATCCACC	60
Oy	1	ATGGATTATGGCATTCACCTGCAAAATGGTATCAAGTCAAAACAGGTGTTAATCCACC	60
Db	1	ATGGATTATGGCATTCACCTGCAAAATGGTATCAAGTCAAAACAGGTGTTAATCCACC	60
Oy	61	GTTGGCGCTGTTGTATGTTAATGAAGATAGATGTTGGTATGTCACACTTATAGAAA	120
Db	61	GTTGGCGCTGTTGTATGTTAATGAAGATAGATGTTGGTATGTCACACTTATAGAAA	120
Oy	121	GGTGCACAGCATGCGGAGGTTCAAGCACTTATATGCGACAAACAAATGCTGAGAGTGGC	180
Db	121	GGTGCACAGCATGCGGAGGTTCAAGCACTTATATGCGACAAACAAATGCTGAGAGTGGC	180
Oy	181	ACGATTATATTTACGTTAGAGCCATGTATGTCATTTTGGTTCACACCACTCTGTGTTAC	240
Db	181	ACGATTATATTTACGTTAGAGCCATGTATGTCATTTTGGTTCACACCACTCTGTGTTAC	240
Oy	241	AAAATTATTCATTGATTAAGATAGCAAAAGTAGTATACCGCAACAAAGACAATTCGTTAGAC	300
Db	241	AAAATTATTCATTGATTAAGATAGCAAAAGTAGTATACCGCAACAAAGACAATTCGTTAGAC	300
Oy	301	ACACATGCTATGAGACGTTACGGGCTACCGGATTTGAGGTGTAATGCGTTGATGATGAA	360
Db	301	ACACATGCTATGAGACGTTACGGGCTACCGGATTTGAGGTGTAATGCGTTGATGATGAA	360

OY	361	CGGGCATCACAATTATACCAAGACCTTTTAAAGCAAAAGCAACCACTGGCAAAATT	420
Dp	361	CGGGCATCACAATTATACCAAGACCTTTTAAAGCAAAAGCAACCACTGGCAAAATT	420
OY	421	ACAGTAAAGATATCTGCAGTTTAGATGCTTAACAACGAGATGATATGACAAAGTCAA	480
Dp	421	ACAGTAAAGATATCTGCAGTTTAGATGCTTAACAACGAGATGATATGACAAAGTCAA	480
OY	481	TGATTTCTTACCAAGAGGTTTAACAAGATGTCCTATTAATTAAAGCATCGACAGACGCA	540
Dp	481	TGATTTCTTACCAAGAGGTTTAACAAGATGTCCTATTAAGTTAAAGCATCGACAGACGCA	540
OY	541	GTTGTTAACTGGAAGAGCTAGACGTTGAAATAGATGATCCACAATATACACAGTATTCAA	600
Dp	541	GTTGTTAACTGGAAGAGCTAGACGTTGAAATAGATGATCCACAATATACACAGTATTCAA	600
OY	601	GATGGAAGAAAACCCCTATTAAGATATATGTCGTAACTGGGAATATTCATTTTAATCAG	660
Dp	601	GATGGAAGAAAACCCCTATTAAGATATATGTCGTAACTGGGAATATTCATTTTAATCAG	660
OY	661	CAAAATTTATCAAGATGATCAACCAACCAATTTGGATATATACTGAAAAATCCAAATTTACA	720
Dp	661	CAAAATTTATCAAGATGATCAACCAACCAATTTGGATATATACTGAAAAATCCAAATTTACA	720
OY	721	AGCAATCAAAACATATTTGAATTTATTACTTGAAGTCTGTGGATTTAAACAACATTTCTT	780
Dp	721	AGCAATCAAAACATATTTGAATTTATTACTTGAAGTCTGTGGATTTAAACAACATTTCTT	780
OY	781	CACATTTATATAAAGAGAGAGTTGGAACTTGTGCTAGTGGAGCAGATCCAAACCTACT	840
Dp	781	CACATTTATATATAAAGAGAGAGTTGGAACTTGTGCTAGTGGAGCAGATCCAAACCTACT	840
OY	841	TCGAAATTCGCATCTATATATATAGATGAATTTATTCTCTATATATGCCCCGAATTAAT	900
Dp	841	TCGAAATTCGCATCTATATATATAGATGAATTTATTCTCTATATATGCCCCGAATTAAT	900
OY	901	GGCGGATCTGGAATATATCAATTTTATCAAAACAANAATGATGATGAGATACACAGATCG	960
Dp	901	GGCGGATCTGGAATATATCAATTTTATCAAAACAANAATGATGATGAGATACACAGATCG	960
OY	961	AACCAATTTGAAATTTGTTCAATTCGAGTTATTTAAATCAAAATGTTAAATTTAACTTTACA	1020
Dp	961	AACCAATTTGAAATTTGTTCAATTCGAGTTATTTAAATCAAAATGTTAAATTTAACTTTACA	1020
OY	1021	AAGAGTGA 1029	
Dp	1021	AAGAGTGA 1029	

```

RESULT 3
PCT-US98-24857-7
: Sequence 7, Application PC/TUS9824857A
: GENERAL INFORMATION:
: APPLICANT: SmithKline Beecham Corporation
: TITLE OF INVENTION: Novel ribG
: FILE REFERENCE: P5044-9
: CURRENT APPLICATION NUMBER: PCT/US98/24857A
: CURRENT FILING DATE: 1998-11-23
: EARLIER APPLICATION NUMBER: 08/978,456
: EARLIER FILING DATE: 1997-11-25
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 7
: LENGTH: 3336
: TYPE: DNA
: ORGANISM: Staphylococcus aureus
PCT-US98-24857-7

```

Query Match	100.0%;	Score 1029;	DB 1;	Length 3336;
Best Local Similarity	100.0%;	Pred. No. 6e-224;		
Matches 1029;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

OY	1	ATGATTTATGCGATTTCACACTTCCAAATATGCTATCAAGCTCAAAACAGGCTTTATTCACACC	60
Db	1	atgattatbgatctcaactctgcaaatatgctaaagctcaaacaggtttaatccacc	60
OY	61	GTTGGCCGTGTTAGTAAATGAAGTAGATGTTGGTATGCTGTCACACTTGAGAAA	120
Db	61	gttggccgtcttgtaagtaataagatagatgttggatgttgctgacaccttgagaaa	120
OY	121	GGTCACAACAGCATCGCGAGGTTCAAGCACTTGATATGCGACAAACAAATGCTGAAGTSCG	180
Db	121	ggtcacaagcctgcggaggttcaaacacttgatatgccaacaacaaatgctgaagttgcg	180
OY	181	ACGATTTATATTACGTGTAAGGCCATGAGTCATTTGGTTCAACACCAACCCTGTTTAC	240
Db	181	acgatttatatactgcttaagagccatgtatgcatcttggctcaaccacccactgtgttaac	240
OY	241	AAATATTGATTGTTAGATAGCAAAAAGTATACGACAAACAAAGACAAATTCTGTACAC	300
Db	241	aaatatttgctgttaagatagcaaaaagtatacgcacaacaaagacaattctgttagac	300
OY	301	ACACATGCTATGACGTTTACGGGCTCACGGGATGAGCTGATGGTGTGATGATGAA	360
Db	301	acacatgctatgtagacgcttaccggctcaccggatctgaagttcgaatgcgttgaatgaa	360
OY	361	CGGCGATCACAAATTATACCAAGACTTTTTTAAAGCAAAAGCAACCAACTGCCAAATT	420
Db	361	cggcgatcaccaattataccaagactctttaaagcaaaaagcaaacagcaactgcacaaatt	420
OY	421	ACAGTGAAGATATCTGCAAGTTTATGATGGTAAACAGCAGATGATATATGACAAAGTAA	480
Db	421	acagtgaagatattctgcaagtttatgattgtaataacaaagcaatgataatggaacaaagtcaa	480
OY	481	TGCAATTCTAACAAGAGGTTTAAACAAGATGCTTAAACTTAAGCAATCGACACAGCCA	540
Db	481	tgcattctactaaacaagggttaaacaaagatgcttaaaagttaagacatcgcacaagcga	540
OY	541	GTTGTTAACTGGAGACGCTACAGTTGAATTAGATGATCCAAATATATCTACACGATTTCAA	600
Db	541	gttgttaacttggaaagctgacagttgattagatgatacccaatatatacgaagttactaa	600
OY	601	GATGGAAAAAACCCTATTAAGAAAGTAAATTTGCTTAAGTGGGAATTAATCATTTTAATCAG	660
Db	601	gatggaaaaaacccctataaagaataatgtctcaagtcgggaatactatcatcttaaccag	660
OY	661	CAAAATTTATCAAGATGATCAACACCAATTTGGATTTACTGAATAATCCAAATTTAACA	720
Db	661	caaatattatacaagtgatcaacaaccaatttggatatatactcgaaatoccaaatttaaca	720
OY	721	AGCAATCAACACATATTGAAATTTATTACTTAAATCTTGTGATTTAACAACATTTCTT	780
Db	721	agcatccaacaacatatttgaatattacttgaagctcttgatttaacaacaactctt	780
OY	781	CACATTTATATAAAGAGAGGTTGGAATTTGCTAGTGGAGCAGGTCAACACCTACT	840
Db	781	cacaattatataaagaagaagttggaactcttgctagtcgagcgcggtccaaaccaactact	840
OY	841	TCAGATTTCTCATCTATTTATATAGATGAATTTATTCTCTATTAATGCCCCGAATTAATT	900
Db	841	tcagatattctccatctatataatagatgaattatctctatatactgccccgaatlaatt	900
OY	901	GGCGATCTCGAATTTATCAATTTTATCAACAAGATGTAAGTGAATACCAATCG	960
Db	901	ggcggaatctcgaaatattcaatattcaacaacaagaatgataatggaagttacagaatgcg	960
OY	961	AACCAATTTGAAATGTTTCATCTCCGATGTTTAAATCAAAAGTTAAATTAATTACTTACA	1020
Db	961	aaccaatttgaaatgttcatctccgagttataatcaacaagaatgtaaatattactttaga	1020
OY	1021	AAGAAGTCA 1029	
Db	1021	aagaagtga 1029	

```

RESULT      4
PCT-US98-25068-7
; Sequence 7, Application PC/TUS9825068
; GENERAL INFORMATION:
; APPLICANT: SmithKline Beecham Corporation
; TITLE OF INVENTION: riba
; FILE REFERENCE: P50444-07
; CURRENT APPLICATION NUMBER: PCT/US98/25068
; CURRENT FILING DATE: 1998-11-24
; EARLIER APPLICATION NUMBER: 08/977,554
; EARLIER FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 3336
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
PCT-US98-25068-7

```

[illegible]

361 CGGCGATCACAATATATACCAAGCTTTTAAAGCAAAAGCAACCACTGCCACAATT 420
421 ACAGTGAATATCTGCAACTTTAGATGTAACCAACGATGATATGACAAAGTCAA 480
421 ACAGTGAATATCTGCAACTTTAGATGTAACCAACGATGATATGACAAAGTCAA 480
481 TGGATTACTACAAAGAGTTAAACAAGATGTCATTAAGTTAAGACATGCACAGACGA 540
481 TGGATTACTACAAAGAGTTAAACAAGATGTCATTAAGTTAAGACATGCACAGACGA 540
541 GTGTTAACTGGAAGACGTACAGTTGAATTAAGATGATCCACAATATCTACACGTATCAA 600
541 GTGTTAACTGGAAGACGTACAGTTGAATTAAGATGATCCACAATATCTACACGTATCAA 600
601 GATGGAAGAAACCTATATAAGTAATATGCTAGCTGGAATATCTCAATTTAATCAG 660
601 GATGGAAGAAACCTATATAAGTAATATGCTAGCTGGAATATCTCAATTTAATCAG 660
661 CAATTTATCAGATGATCAACACCAATTTGATATATCTGAAATCCAAATTTTACA 720
661 CAATTTATCAGATGATCAACACCAATTTGATATATCTGAAATCCAAATTTTACA 720
721 AGCAATCAACACATATTTGAATTTATTTACTGTAAGTCTTGATTTAACAACATTTCTT 780
721 AGCAATCAACACATATTTGAATTTATTTACTGTAAGTCTTGATTTAACAACATTTCTT 780
781 CACATTTATATAAAGAGAGTGGAACTTGTCTAGTCCAGCAGGTCCACACCTACT 840
781 CACATTTATATAAAGAGAGTGGAACTTGTCTAGTCCAGCAGGTCCACACCTACT 840
841 TCAGATTTCCATCTATTTATATAGATGATTTATTTCTATTTATGCGCGAAATTTAT 900
841 TCAGATTTCCATCTATTTATATAGATGATTTATTTCTATTTATGCGCGAAATTTAT 900
901 GCGGATCTGGAATTAATCAATTTTATCAACAACATGATGATGATGATGATGATGCG 960
901 GCGGATCTGGAATTAATCAATTTTATCAACAACATGATGATGATGATGATGATGCG 960
961 AACCAATTTGAATTTGTCATTTCCGATTTATTAATCAAAATGTTAATTAATTTACGA 1020
961 AACCAATTTGAATTTGTCATTTCCGATTTATTAATCAAAATGTTAATTAATTTACGA 1020
1021 AAGAGTGA 1029
1021 AAGAGTGA 1029

ATTORNEY/AGENT INFORMATION:
NAME: Dickinson, Todd O
REGISTRATION NUMBER: 28,354
REFERENCE/DOCKET NUMBER: P50444-8
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-994-2252
TELEFAX: 215-994-2222
TELEX:
INFORMATION FOR SEO ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3336 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
us-08-977-866-7
Query Match 100.0%; Score 1029; DB 32; Length 3336;
Best Local Similarity 100.0%; Pred. No. 6e-224;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 ATGATTTATGCGATTCGAACTGCAATATGTTGTCAGAGTCAACACAGGTGTAATCCACC 60
1 ATGATTTATGCGATTCGAACTGCAATATGTTGTCAGAGTCAACACAGGTGTAATCCACC 60
61 GTTGGCGCTGTGTAGTTAATGAAGTATGTTGTTATGTTGTCACACTGAGAAAA 120
61 GTTGGCGCTGTGTAGTTAATGAAGTATGTTGTTATGTTGTCACACTGAGAAAA 120
121 GGTGACAGCATCGGAGGTTCAAGCATTGATATGTCACACAAATATCTGAGGTGG 180
121 GGTGACAGCATCGGAGGTTCAAGCATTGATATGTCACACAAATATCTGAGGTGG 180
121 GGTGACAGCATCGGAGGTTCAAGCATTGATATGTCACACAAATATCTGAGGTGG 180
181 ACATTTATTTACGTTAGAGCCATGATCATTTTGTGTTCAACACCACCTGTGTTAAC 240
181 ACATTTATTTACGTTAGAGCCATGATCATTTTGTGTTCAACACCACCTGTGTTAAC 240
181 ACATTTATTTACGTTAGAGCCATGATCATTTTGTGTTCAACACCACCTGTGTTAAC 240
241 AAAATTTATGATTTGAATAGCAAAAGTATGATACGCAACAAAGCAATTTGTTAGAC 300
241 AAAATTTATGATTTGAATAGCAAAAGTATGATACGCAACAAAGCAATTTGTTAGAC 300
301 ACATGATGATGATGAGAGTTCACGGCTCAAGGTTATGAGGTTGAAATGCGTTGATGAA 360
301 ACATGATGATGATGAGAGTTCACGGCTCAAGGTTATGAGGTTGAAATGCGTTGATGAA 360
301 ACATGATGATGATGAGAGTTCACGGCTCAAGGTTATGAGGTTGAAATGCGTTGATGAA 360
361 CGGCGATCACAATTAATACCAAGCTTTTAAAGCAAAAGCAAAAGCAAGTGCACAAATT 420
361 CGGCGATCACAATTAATACCAAGCTTTTAAAGCAAAAGCAAAAGCAAGTGCACAAATT 420
421 ACAGTGAAGTATCTGCAAGTTAGATGTAACCAAGCGAATGATATAGACAAAGTCAA 480
421 ACAGTGAAGTATCTGCAAGTTAGATGTAACCAAGCGAATGATATAGACAAAGTCAA 480
481 TGGATTACTACAAAGAGTTAAACAAGATGTCATTAAGTTAAGACATGCACAGACGA 540
481 TGGATTACTACAAAGAGTTAAACAAGATGTCATTAAGTTAAGACATGCACAGACGA 540
541 GTGTTAACTGGAAGACGTACAGTTGAATTAAGATGATCCACAATATCTACACGTATCAA 600
541 GTGTTAACTGGAAGACGTACAGTTGAATTAAGATGATCCACAATATCTACACGTATCAA 600
601 GATGGAAGAAACCTATATAAGTAATATGCTAGCTGGAATATCTCAATTTAATCAG 660
601 GATGGAAGAAACCTATATAAGTAATATGCTAGCTGGAATATCTCAATTTAATCAG 660
661 CAATTTATCAGATGATCAACACCAATTTGATATATCTGAAATCCAAATTTTACA 720
661 CAATTTATCAGATGATCAACACCAATTTGATATATCTGAAATCCAAATTTTACA 720
721 AGCAATCAACACATATTTGAATTTATTTACTGTAAGTCTTGATTTAACAACATTTCTT 780
721 AGCAATCAACACATATTTGAATTTATTTACTGTAAGTCTTGATTTAACAACATTTCTT 780

STREET: 4000 Bell Atlantic Tower, 1717 Arch Street
CITY: Philadelphia
STATE: PA
COUNTRY: US
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEO for Windows Version 2.0
CURRENT APPLICATION DATA: 479
APPLICATION NUMBER: US/09/492.479
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/977,866
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dickinson, Todd O
REGISTRATION NUMBER: 28,354
REFERENCE/DOCKET NUMBER: P50444-8
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-994-2252
TELEFAX: 215-994-2222
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3336 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-492-479-7

Query Match	100.0%;	Score 1029;	DB 54;	Length 3336;
Best Local Similarity	100.0%;	Pred. No. 6e-224;		
Matches 1029;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0.

QY	1	ATGATTATATGGCATTCACACTGCAAAATATGGTACAAAGGTCAAAACAGGTGTTATATCCACC	60
Db	1	ATGGATTATATGGGATTCACACTTGCAAAATATGGTACAAAGGTCAAAACAGGTGTTATATCCACC	60
QY	61	GTTGGCGCTGTTGTAGTTAATGAAGGTAGGATCTGTGGTATGGTGCAACTTGAGAAA	120
Db	61	GTTGGCGCTGTTGTAGTTAATGAAGGTAGGATCTGTGGTATGGTGCAACTTGAGAAA	120
QY	121	GGTGACAAAGCATGCGGAGGTTCAAGCACTTATATGGCACAAACAAATGCTGAAGTGGC	180
Db	121	GGTGACAAAGCATGCGGAGGTTCAAGCACTTATATGGCACAAACAAATGCTGAAGTGGC	180
QY	181	ACGATTATATTTACGTTAGAGCCATGTAGTCATTTTGGTTCACACCAACCCTGTGTTAAC	240
Db	181	ACGATTATATTTACGTTAGAGCCATGTAGTCATTTTGGTTCACACCAACCCTGTGTTAAC	240
QY	241	AAATTTATGATTTCTAGATAGCAAAAGTACGTATACGCAACAAAGCAATTCGTTAGAC	300
Db	241	AAATTTATGATTTCTAGATAGCAAAAGTACGTATACGCAACAAAGCAATTCGTTAGAC	300
QY	301	ACACATGGTGTATGAGACGTTACGGGCTCACCGTATTGAGTTGTAATCGTTGATGATGAA	360
Db	301	ACACATGGTGTATGAGACGTTACGGGCTCACCGTATTGAGTTGTAATCGTTGATGATGAA	360
QY	361	CGGGCATCACAAATATATACCAAGACTTTTATTAAGCAAAAGCAAAAGCAACTGCCCAAAAT	420
Db	361	CGGGCATCACAAATATATACCAAGACTTTTATTAAGCAAAAGCAAAAGCAACTGCCCAAAAT	420
QY	421	ACAGTGAAGATATGTCAGAGTTTACATGCTTAACCAAGCGAATGATAATGGCAAAAGTCAA	480
Db	421	ACAGTGAAGATATGTCAGAGTTTACATGCTTAACCAAGCGAATGATAATGGCAAAAGTCAA	480
QY	481	TGGATTACTAACCAAGAGGTTAAACCAAGATGTCATTAAGTAAAGACATCGACACGACGCA	540
Db	481	TGGATTACTAACCAAGAGGTTAAACCAAGATGTCATTAAGTAAAGACATCGACACGACGCA	540

Oy	541	GTGTTAACTGGAAGACCTAACAGTTGAATTAGATGATCCACATATATACAGCTTTCAA	600
Dd	541	GTGTTAACTGGAAGACCTAACAGTTGAATTAGATGATCCACATATATACAGCTTTCAA	600
Oy	601	GATGGAAAAAACCCCTATTTAAAGTAAATATTTGCTAAAGTCGGGAATATTCATTATATCAG	660
Dd	601	GATGGAAAAAACCCCTATTTAAAGTAAATATTTGCTAAAGTCGGGAATATTCATTATATCAG	660
Oy	661	CAAAATTTATCAAGATGATCAACCAACCAATTTGGATATATACTGAAAAATCCAAATTTAA	720
Dd	661	CAAAATTTATCAAGATGATCAACCAACCAATTTGGATATATACGAAAAATCCAAATTTAA	720
Oy	721	AGCAATCAACACATATTTGAATATTTTACTGTGAAGCTGTGGATTTAAACAATTCCT	780
Dd	721	AGCAATCAACACATATTTGAATATTTTACTGTGAAGCTGTGGATTTAAACAATTCCT	780
Oy	781	CACATTTATTTAAAGAAGTTGGAACCTTGCTAGTCGAGCGCAACCACTACT	840
Dd	781	CACATTTATTTAAAGAAGTTGGAACCTTGCTAGTCGAGCGCAACCACTACT	840
Oy	841	TCAGATTCCTCACTCTTTATATAGATGAATTTATCTCTATTTATGCCCGGAATTAAT	900
Dd	841	TCAGATTCCTCACTCTTTATATAGATGAATTTATCTCTATTTATGCCCGGAATTAAT	900
Oy	901	GCGGGAATCTGGAATTTATCAATTTTATCAACAACAATGATGATTGAGATACCAAGATCG	960
Dd	901	GCGGGAATCTGGAATTTATCAATTTTATCAACAACAATGATGATTGAGATACCAAGATCG	960
Oy	961	AACCAATTTGAATTTGTTCAATTTCCGAGTTATTAATCAAAATGTTTAATTAACCTTACGA	1020
Dd	961	AACCAATTTGAATTTGTTCAATTTCCGAGTTATTAATCAAAATGTTTAATTAACCTTACGA	1020
Oy	1021	AAGAAGTGA 1029	
Dd	1021	AAGAAGTGA 1029	

RESULT 12
US-09-376-633-7
; Sequence 7, Application US/09376633
; GENERAL INFORMATION:
; APPLICANT: Palmer, Leslie M.
; TITLE OF INVENTION: Novel rDNA
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Street
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/376,633
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/978,456
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, Todd O
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: P50444-9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2252
; TELEFAX: 215-994-2222
; TELETYPE:
; INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:
LENGTH: 3336 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-376-633-7

Query Match 100.0%; Score 1029; DB 103; Length 3336;
Best Local Similarity 100.0%; Pred. No. 6e-224;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGATATTCGATTCACCTTGCACCAATATGTACACAGTCAACAGGTGTTATCCACC 60
    |||||||
DB 1 ATGATATTCGATTCACCTTGCACCAATATGTACACAGTCAACAGGTGTTATCCACC 60

QY 61 GTTGGCGCTGTTAGTTAAATGAAGTAGAGATTGTTGTTGTCACACTTGAGAAA 120
    |||||||
DB 61 GTTGGCGCTGTTAGTTAAATGAAGTAGAGATTGTTGTTGTCACACTTGAGAAA 120

QY 121 GGTGACACAGATGGGAGGTTCAAGCACTTGATGTCACACAAATGCTGAAGGTGCG 180
    |||||||
DB 121 GGTGACACAGATGGGAGGTTCAAGCACTTGATGTCACACAAATGCTGAAGGTGCG 180

QY 181 ACGATTATATACGTAGAGCCATGTAGTCAATTTGTTCAACACACCCTGTGTAAAC 240
    |||||||
DB 181 ACGATTATATACGTAGAGCCATGTAGTCAATTTGTTCAACACACCCTGTGTAAAC 240

QY 241 AAAATTATGATTGTAAGATAGCAAAAGTAGATACGCCAACAAAGCAATTCGTTAGAC 300
    |||||||
DB 241 AAAATTATGATTGTAAGATAGCAAAAGTAGATACGCCAACAAAGCAATTCGTTAGAC 300

QY 301 ACACATGCTATGAGAGTACCGGCTCAGCGTATGAGTGAATGCGTTGATGATGAA 360
    |||||||
DB 301 ACACATGCTATGAGAGTACCGGCTCAGCGTATGAGTGAATGCGTTGATGATGAA 360

QY 361 CGGGCATCACAAATTAATCCAAAGACTTTTAAAGCAAAAGCAACACTGCCAATTT 420
    |||||||
DB 361 CGGGCATCACAAATTAATCCAAAGACTTTTAAAGCAAAAGCAACACTGCCAATTT 420

QY 421 ACAAGTAAGTATCTGCAAGTTAGATGTAAACACGCAATGATTAATGCAAAAGTCAA 480
    |||||||
DB 421 ACAAGTAAGTATCTGCAAGTTAGATGTAAACACGCAATGATTAATGCAAAAGTCAA 480

QY 481 TGGATTCTACAAAGAGGTTAAACCAAGATGCTATTAAGTTAAACATCGACAGAGCA 540
    |||||||
DB 481 TGGATTCTACAAAGAGGTTAAACCAAGATGCTATTAAGTTAAACATCGACAGAGCA 540

QY 541 GTGTTACTGGAAGACGTACAGTTGAATAGATGCACAAATATCTACACGATTTCAA 600
    |||||||
DB 541 GTGTTACTGGAAGACGTACAGTTGAATAGATGCACAAATATCTACACGATTTCAA 600

QY 601 GATGAAAAAACCTATTAAGTAATATGCTAAGTCTGGGAATATTCATTTTAATCAG 660
    |||||||
DB 601 GATGAAAAAACCTATTAAGTAATATGCTAAGTCTGGGAATATTCATTTTAATCAG 660

QY 661 CAATTTATCAAGTGAATCAACCAATTTGATATATCTGAAAAATCCAAATTTTACA 720
    |||||||
DB 661 CAATTTATCAAGTGAATCAACCAATTTGATATATCTGAAAAATCCAAATTTTACA 720

QY 721 AGCAATCAACATATGTAATATTTACTGTAAGTCTGTGATTAACAAATCTT 780
    |||||||
DB 721 AGCAATCAACATATGTAATATTTACTGTAAGTCTGTGATTAACAAATCTT 780

QY 781 CACAATTTATTAAGAGAGGTGGAAGTCTGAGGAGGTCACCACTACT 840
    |||||||
DB 781 CACAATTTATTAAGAGAGGTGGAAGTCTGAGGAGGTCACCACTACT 840

QY 841 TCAGAAATTCATCTATTTATATGATGAATTTATCTCTATTTAGCCCCGAATTAAT 900
    |||||||
DB 841 TCAGAAATTCATCTATTTATATGATGAATTTATCTCTATTTAGCCCCGAATTAAT 900

QY 901 GGGGATCTGGAATATTCATTTATCAACAAATGATGATGATGATACCAATGCG 960
    |||||||
```

```
DB 901 GGGGATCTGGAATATTCATTTATCAACAAATGATGATGATGATACCAATGCG 960
    |||||||
QY 961 ACCCAATTTAAATGTTCAATTCGAGTATTTAAATCAAAATGTTAAATTAATCA 1020
    |||||||
DB 961 ACCCAATTTAAATGTTCAATTCGAGTATTTAAATCAAAATGTTAAATTAATCA 1020
    |||||||

QY 1021 AAGAGTGA 1029
    |||||||
DB 1021 AAGAGTGA 1029
```

RESULT 13

```
US-08-827-356-1376
; Sequence 1376 Application US/08827356
; GENERAL INFORMATION:
; APPLICANT: George H. Shimer, Jr.
; APPLICANT: George H. Miller
; APPLICANT: Roberta S. Hare
; APPLICANT: Karen J. Shaw
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS RELATED
; TITLE OF INVENTION: COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 5574
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering-Plough Corporation
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07033-0530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/827,356
; FILING DATE: 01-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/014,477
; FILING DATE: 01-APR-1996
; APPLICATION NUMBER: 60/016,743
; FILING DATE: 02-MAY-1996
; APPLICATION NUMBER: 60/020,016
; FILING DATE: 14-JUN-1996
; INFORMATION FOR SEQ. ID NO: 1376:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1119 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: 1...1119
US-08-827-356-1376
```

Query Match 97.8%; Score 1006.4; DB 23; Length 1119;
Best Local Similarity 99.1%; Pred. No. 6.2e-219;
Matches 1023; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

```
QY 1 ATGATATTCGATTCACCTTGCACCAATATGTACACAGTCAACAGGTGTTATCCACC 60
    |||||||
DB 87 ATGATATTCGATTCACCTTGCACCAATATGTACACAGTCAACAGGTGTTATCCACC 146
    |||||||

QY 61 GTTGGCGCTGTTAGTTAAATGAAGTAGAGATTGTTGTTGTCACACTTGAGAAA 120
    |||||||
DB 147 GTTGGCGCTGTTAGTTAAATGAAGTAGAGATTGTTGTTGTCACACTTGAGAAA 206
    |||||||
```


Db 208 ggtgacaacatgctggaaggttcaagcattgatgacacacaaatgctgaagtgcg 267
OY 181 ACGATTATATTCGTTAGAGCCATGATGATTTGGTTCAACACACCCCTGTTAAC 240
Db 268 acgattatattcgttagagccattgacatttggttcaacacacccctgttaac 327
OY 241 AAAATTATGATTGTAAAGATGCAAAAAGTATACGCAACAAAGCAATTGTTAGAC 300
Db 328 aaattatttgattgtagaagtagcaaaagtattatgcgcaacaaagcaattcgtagac 387
OY 301 ACACATGCTGATGACAGCTTACGGGCTCACGGTATTGAGGTTGAGTGGTATGATGA 360
Db 388 acacatggtgatgagacggttacggtccacgattgaagttgaatcggttgatgaa 447
OY 361 CGGGCATTCACAAATTATACCAAGACTTTTAAAGCAAAAGCAACGCAACAATT 420
Db 448 cgggcatcacattatcaacaaagcctttcaagcaaaagcaagcaacgcaacatt 507
OY 421 ACAGTGAAGTATCTGCAGTTAGATGCTAAACAAAGCGATGTAATGGAACAACTCA 480
Db 508 acagtgaaagtatctgcaagcttagatgtaaacaaagcgaatgataatgcaaaagtaa 567
OY 481 TGATTATCTACAAAGAGTTAAACAAGATGCTATATAGTTAAGACATCGACAGACGCA 540
Db 568 tggattactacaagaaggttaacaagaatgctatagttgaagacatcgacacgacgca 627
OY 541 GTCTTAACGTGAGAGCTAGCTGATTTGATGATGATCCAAATATCTAAGCTATTC 600
Db 628 ggttaactggaagacgttaacagttgaatgatacacaatactacacgtatcaaa 687
OY 601 GATGGAAGAAACCCCTATTAAGTAAATATTTGCTAAGTCTGGGATATTCATTTTAATCAG 660
Db 688 gatggaagaaacccctataaagaatatactgctcaagctcgtggaatatacttttaacag 747
OY 661 CAATTATTCAGATGATGATCAACCAATTTGGATATATCTGAAAACTCAATTTTACA 720
Db 748 caatttatcaagatgaatcaacacacatttgatatactgaaaaatccaaatttaca 807
OY 721 AGCAATCAACACATATTTGAAATTTATTTTACTTGAAGCTTGATTTAAACAATCTT 780
Db 808 agaatcaaacacatttgaattatcttgaagctcgttgatatttaaaacaattcct 867
OY 781 CACAATTTATATAAGAGAGAGTGGAACTTCTAGTCGAGCAGGTCACACACTACT 840
Db 868 caaatattatataaagagagtgtagaacttgtagtcgaagcaggtccaaacacact 927
OY 841 TCAGAAATTCCTCATCT--ATTATATGATGAATTTATTTCTATTTATGCCCCGAATTA 897
Db 928 tcgaatttctccaatctaatatataatagaaatttattctctattatgccccgaattta 987
OY 898 ATTGGCGGATCTGGAATTTATCAATTTTATCAACAAATGATGATGATGATACCAAT 957
Db 988 attgagcgagactggaatttcaatttcaacaaatgtagatgtagataccagat 1047
OY 958 GCGAACCAATTTGAATTTGTTCAATTCGAGTTATTAATTAATTAATTAATTAATTTA 1017
Db 1048 ggcgaacaaattgtaaatgttccatccgagttatataatcaaatgttaataactta 1107
OY 1018 CGAAGAGAGTGA 1029
Db 1108 cgaagaagagtgta 1119

RESULT 15
US-09-611-529-7446/c
; Sequence 7446, Application US/09611529
; GENERAL INFORMATION:
; APPLICANT: George H. Shimer, Jr.
; APPLICANT: George H. Shimer, Jr.
; APPLICANT: Robert S. Hare
; APPLICANT: Karen J. Shaw
; TITLE OF INVENTION: Staphylococcus aureus Related Compositions and Methods
; FILE REFERENCE: 1034/IC963US1

; CURRENT APPLICATION NUMBER: US/09/611,529
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/417,811
; PRIOR FILING DATE: 1999-10-14
; PRIOR APPLICATION NUMBER: US 09/353,718
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/266,557
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,556
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,555
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,542
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,541
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/037,934
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: US 09/036,720
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,338
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,334
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,221
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,137
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,082
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,081
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,079
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/035,913
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/035,744
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 08/827,356
; PRIOR FILING DATE: 1997-04-01
; PRIOR APPLICATION NUMBER: US 08/831,156
; PRIOR FILING DATE: 1997-04-01
; PRIOR APPLICATION NUMBER: US 60/014,477
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: US 60/016,743
; PRIOR FILING DATE: 1996-05-02
; PRIOR APPLICATION NUMBER: US 60/020,016
; PRIOR FILING DATE: 1996-06-14
; NUMBER OF SEQ ID NOS: 7451
; SEQ ID NO 7446
; LENGTH: 74105
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-611-529-7446

Query Match 97.8%; Score 1006.4; DB 57; Length 74105;
Best Local Similarity 99.1%; Pred. No. 2e-218;
Matches 1023; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

OY 1 ATGATTATGCGATTCACTTGCAGTAATATGTAACAAGTCAACAGGTGTTATCCACC 60
Db 24245 ATGATTATGCGATTCACTTGCAGTAATATGTAACAAGTCAACAGGTGTTATCCACC 24186
OY 61 GTTGGCGCTGTTGTTACTTAATGAGGTAGAGATTGTTGTTATGTTGTCACACTTGAGAAAA 120
Db 24185 GTTGGCGCTGTTGTTACTTAATGAGGTAGAGATTGTTGTTATGTTGTCACACTTGAGAAAA 24126
OY 121 GGTGACAAGATCGGAGGTCAAGCACTTGATATGCAACCAACAAATGCTGAAGTGGC 180
Db 24125 GGTGACAAGATCGGAGGTCAAGCACTTGATATGCAACCAACAAATGCTGAAGTGGC 24066
OY 181 ACGATTATATTCGTTAGAGCCATGATGATTTGGTTCAACACACCCCTGTGTTAAC 240

```
|||||
Db 24065 ACGATTTATATTAGCTTAGACCAATGATGTTGGTTCAACACCCCTGTGTTAAC 24006
OY 241 AAAATTTATGTTGTAAGATAGCAAAAGTATGATGACAAACAAATTCGTTAGAC 300
Db 24005 AAAATTTATGTTGTAAGATAGCAAAAGTATGATGACAAACAAATTCGTTAGAC 23946
OY 301 ACACATGCTGATGAGACGCTTACGGGCTCAGGTTATGAGTTGAATGCGTTGATGATA 360
Db 23945 ACACATGCTGATGAGACGCTTACGGGCTCAGGTTATGAGTTGAATGCGTTGATGATA 23886
OY 361 CGGCATCACAATTATACCAAGACTTTTAAAGCAAAAGCAAACTGCCACAATT 420
Db 23885 CGGCATCACAATTATACCAAGACTTTTAAAGCAAAAGCAAACTGCCACAATT 23826
OY 421 ACAGTGAAGTATCTGCAAGTTTGAATGTAACAAAGCAATGATTAATGACAAAGTCAA 480
Db 23825 ACAGTGAAGTATCTGCAAGTTTGAATGTAACAAAGCAATGATTAATGACAAAGTCAA 23766
OY 481 TGGATTACTTAACAAAGAGTTAAACAGATGCTATAAGTTAAGACATGACACGACGCA 540
Db 23765 TGGATTACTTAACAAAGAGTTAAACAGATGCTATAAGTTAAGACATGACACGACGCA 23706
OY 541 GTGTTAACTGGAAGACGTAAGTTGAATGATGATCCACAATATACACGTAATCAA 600
Db 23705 GTGTTAACTGGAAGACGTAAGTTGAATGATGATCCACAATATACACGTAATCAA 23646
OY 601 GATGGAAGAAACCCGATTAAGTAAATATTGCTAAGCTGGGAATATTCATTTTAACTAG 660
Db 23645 GATGGAAGAAACCCGATTAAGTAAATATTGCTAAGCTGGGAATATTCATTTTAACTAG 23586
OY 661 CAAATTTATCAAGATGAATCAACACCAATTTGGATATATCTGAAATCCAAATTTAACA 720
Db 23585 CAAATTTATCAAGATGAATCAACACCAATTTGGATATATCTGAAATCCAAATTTAACA 23526
OY 721 AGCAATCAACACATATTTGAATTTTACTTGAAGCTTGTGATTTAACAACAATTCCT 780
Db 23525 AGCAATCAACACATATTTGAATTTTACTTGAAGCTTGTGATTTAACAACAATTCCT 23466
OY 781 CACAATTTATTAAGAGAGGAGTTGGAATTTGCTAGTCGAGGAGGTCACCAACCTACT 840
Db 23465 CACAATTTATTAAGAGAGGAGTTGGAATTTGCTAGTCGAGGAGGTCACCAACCTACT 23406
OY 841 TCAGAAATTCCTCATCT--ATTATATAGATGAATTTATCTCTATTATGCCCGAAATTA 897
Db 23405 TCAGAAATTCCTCATCTAATATATATAGATGAATTTATCTCTATTATGCCCGAAATTA 23346
OY 898 ATTGGCGGATCTGCAAAATTTATCAATTTATCAACAAATGATGATGAGTACCAGAT 957
Db 23345 ATTGGCGGATCTGCAAAATTTATCAATTTATCAACAAATGATGATGAGTACCAGAT 23286
OY 958 GCGAACCATTGTAATGTTCAATCGAGTTATTAATCAAAATGTTAAATTAACCTTA 1017
Db 23285 GCGAACCATTGTAATGTTCAATCGAGTTATTAATCAAAATGTTAAATTAACCTTA 23226
OY 1018 CGAAAGAGTGA 1029
Db 23225 CGAAAGAGTGA 23214
```

Search completed: November 19, 2000, 05:56:05
Job time: 8392 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 19, 2000, 05:56:05 : Search time 2066.03 Seconds
(without alignments)
1263.820 Million cell updates/sec

Title: US-08-978-456-3
Perfect score: 1269
Sequence: 1 AANCACCAATCCNATGGCA.....GTTAACACAGATGCTATAG 1269

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10041192 seqs, 1028798125 residues

Total number of hits satisfying chosen parameters: 20082384

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending_Patents_NA:*
1: /cgn2_6/ptodata/2/pna/US06.COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06.COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07.COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US08.COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US08A.COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US08B.COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US08C.COMB.seq:*
8: /cgn2_6/ptodata/2/pna/US08D.COMB.seq:*
9: /cgn2_6/ptodata/2/pna/US08E.COMB.seq:*
10: /cgn2_6/ptodata/2/pna/US08F.COMB.seq:*
11: /cgn2_6/ptodata/2/pna/US08G.COMB.seq:*
12: /cgn2_6/ptodata/2/pna/US08H.COMB.seq:*
13: /cgn2_6/ptodata/2/pna/US08I.COMB.seq:*
14: /cgn2_6/ptodata/2/pna/US08J.COMB.seq:*
15: /cgn2_6/ptodata/2/pna/US08K.COMB.seq:*
16: /cgn2_6/ptodata/2/pna/US08L.COMB.seq:*
17: /cgn2_6/ptodata/2/pna/US08M.COMB.seq:*
18: /cgn2_6/ptodata/2/pna/US08N.COMB.seq:*
19: /cgn2_6/ptodata/2/pna/US08O.COMB.seq:*
20: /cgn2_6/ptodata/2/pna/US08P.COMB.seq:*
21: /cgn2_6/ptodata/2/pna/US08Q.COMB.seq:*
22: /cgn2_6/ptodata/2/pna/US08R.COMB.seq:*
23: /cgn2_6/ptodata/2/pna/US08S.COMB.seq:*
24: /cgn2_6/ptodata/2/pna/US08T.COMB.seq:*
25: /cgn2_6/ptodata/2/pna/US08U.COMB.seq:*
26: /cgn2_6/ptodata/2/pna/US08V.COMB.seq:*
27: /cgn2_6/ptodata/2/pna/US08W.COMB.seq:*
28: /cgn2_6/ptodata/2/pna/US08X.COMB.seq:*
29: /cgn2_6/ptodata/2/pna/US08Y.COMB.seq:*
30: /cgn2_6/ptodata/2/pna/US08Z.COMB.seq:*
31: /cgn2_6/ptodata/2/pna/US09A.COMB.seq:*
32: /cgn2_6/ptodata/2/pna/US09B.COMB.seq:*
33: /cgn2_6/ptodata/2/pna/US09C.COMB.seq:*
34: /cgn2_6/ptodata/2/pna/US09D.COMB.seq:*
35: /cgn2_6/ptodata/2/pna/US09E.COMB.seq:*
36: /cgn2_6/ptodata/2/pna/US09F.COMB.seq:*
37: /cgn2_6/ptodata/2/pna/US09G.COMB.seq:*
38: /cgn2_6/ptodata/2/pna/US09H.COMB.seq:*
39: /cgn2_6/ptodata/2/pna/US09I.COMB.seq:*
40: /cgn2_6/ptodata/2/pna/US09J.COMB.seq:*
41: /cgn2_6/ptodata/2/pna/US09K.COMB.seq:*
42: /cgn2_6/ptodata/2/pna/US09L.COMB.seq:*
43: /cgn2_6/ptodata/2/pna/US09M.COMB.seq:*

44: /cgn2_6/ptodata/2/pna/US09N.COMB.seq:*
45: /cgn2_6/ptodata/2/pna/US09O.COMB.seq:*
46: /cgn2_6/ptodata/2/pna/US09P.COMB.seq:*
47: /cgn2_6/ptodata/2/pna/US09Q.COMB.seq:*
48: /cgn2_6/ptodata/2/pna/US09R.COMB.seq:*
49: /cgn2_6/ptodata/2/pna/US09S.COMB.seq:*
50: /cgn2_6/ptodata/2/pna/US09T.COMB.seq:*
51: /cgn2_6/ptodata/2/pna/US09U.COMB.seq:*
52: /cgn2_6/ptodata/2/pna/US09V.COMB.seq:*
53: /cgn2_6/ptodata/2/pna/US09W.COMB.seq:*
54: /cgn2_6/ptodata/2/pna/US09X.COMB.seq:*
55: /cgn2_6/ptodata/2/pna/US09Y.COMB.seq:*
56: /cgn2_6/ptodata/2/pna/US09Z.COMB.seq:*
57: /cgn2_6/ptodata/2/pna/US09A.COMB.seq:*
58: /cgn2_6/ptodata/2/pna/US09B.COMB.seq:*
59: /cgn2_6/ptodata/2/pna/US09C.COMB.seq:*
60: /cgn2_6/ptodata/2/pna/US09D.COMB.seq:*
61: /cgn2_6/ptodata/2/pna/US09E.COMB.seq:*
62: /cgn2_6/ptodata/2/pna/US09F.COMB.seq:*
63: /cgn2_6/ptodata/2/pna/US09G.COMB.seq:*
64: /cgn2_6/ptodata/2/pna/US09H.COMB.seq:*
65: /cgn2_6/ptodata/2/pna/US09I.COMB.seq:*
66: /cgn2_6/ptodata/2/pna/US09J.COMB.seq:*
67: /cgn2_6/ptodata/2/pna/US09K.COMB.seq:*
68: /cgn2_6/ptodata/2/pna/US09L.COMB.seq:*
69: /cgn2_6/ptodata/2/pna/US09M.COMB.seq:*
70: /cgn2_6/ptodata/2/pna/US09N.COMB.seq:*
71: /cgn2_6/ptodata/2/pna/US09O.COMB.seq:*
72: /cgn2_6/ptodata/2/pna/US09P.COMB.seq:*
73: /cgn2_6/ptodata/2/pna/US09Q.COMB.seq:*
74: /cgn2_6/ptodata/2/pna/US09R.COMB.seq:*
75: /cgn2_6/ptodata/2/pna/US09S.COMB.seq:*
76: /cgn2_6/ptodata/2/pna/US09T.COMB.seq:*
77: /cgn2_6/ptodata/2/pna/US09U.COMB.seq:*
78: /cgn2_6/ptodata/2/pna/US09V.COMB.seq:*
79: /cgn2_6/ptodata/2/pna/US09W.COMB.seq:*
80: /cgn2_6/ptodata/2/pna/US09X.COMB.seq:*
81: /cgn2_6/ptodata/2/pna/US09Y.COMB.seq:*
82: /cgn2_6/ptodata/2/pna/US09Z.COMB.seq:*
83: /cgn2_6/ptodata/2/pna/US09A.COMB.seq:*
84: /cgn2_6/ptodata/2/pna/US09B.COMB.seq:*
85: /cgn2_6/ptodata/2/pna/US09C.COMB.seq:*
86: /cgn2_6/ptodata/2/pna/US09D.COMB.seq:*
87: /cgn2_6/ptodata/2/pna/US09E.COMB.seq:*
88: /cgn2_6/ptodata/2/pna/US09F.COMB.seq:*
89: /cgn2_6/ptodata/2/pna/US09G.COMB.seq:*
90: /cgn2_6/ptodata/2/pna/US09H.COMB.seq:*
91: /cgn2_6/ptodata/2/pna/US09I.COMB.seq:*
92: /cgn2_6/ptodata/2/pna/US09J.COMB.seq:*
93: /cgn2_6/ptodata/2/pna/US09K.COMB.seq:*
94: /cgn2_6/ptodata/2/pna/US09L.COMB.seq:*
95: /cgn2_6/ptodata/2/pna/US09M.COMB.seq:*
96: /cgn2_6/ptodata/2/pna/US09N.COMB.seq:*
97: /cgn2_6/ptodata/2/pna/US09O.COMB.seq:*
98: /cgn2_6/ptodata/2/pna/US09P.COMB.seq:*
99: /cgn2_6/ptodata/2/pna/US09Q.COMB.seq:*
100: /cgn2_6/ptodata/2/pna/US09R.COMB.seq:*
101: /cgn2_6/ptodata/2/pna/US09S.COMB.seq:*
102: /cgn2_6/ptodata/2/pna/US09T.COMB.seq:*
103: /cgn2_6/ptodata/2/pna/US09U.COMB.seq:*
104: /cgn2_6/ptodata/2/pna/US09V.COMB.seq:*
105: /cgn2_6/ptodata/2/pna/US09W.COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1250	98.5	1269	1	PCR-US97-02318-946 Sequence 946, App

Sequence 3, Appl1
Sequence 946, App
Sequence 3, Appl
Sequence 7446, Ap
Sequence 107, Ap
Sequence 142, App
Sequence 142, App
Sequence 142, App
Sequence 616, App
Sequence 618, App
Sequence 2079, Ap
Sequence 1376, Ap
Sequence 1, Appl
Sequence 1, Appl
Sequence 7, Appl
Sequence 7, Appl
Sequence 7, Appl
Sequence 5, Appl
Sequence 5, Appl
Sequence 7, Appl
Sequence 7, Appl
Sequence 7, Appl
Sequence 7, Appl
Sequence 7, Appl
Sequence 7, Appl
Sequence 7, Appl
Sequence 710, App
Sequence 711, App
Sequence 1118, Ap
Sequence 1118, Ap
Sequence 1118, Ap
Sequence 1118, Ap
Sequence 2343, Ap
Sequence 2157, Ap
Sequence 2251, Ap
Sequence 1, Appl
Sequence 92, Appl
Sequence 7438, Ap
Sequence 1376, Ap
Sequence 1676, Ap
Sequence 1676, Ap

RESULT 1

Sequence 946 Application PC/TUS9702318
GENERAL INFORMATION:
APPLICANT: Black, Michael
APPLICANT: Burnham, Martin
APPLICANT: Hodgson, John
APPLICANT: Knowles, David
APPLICANT: Nicholas, Richard
APPLICANT: Pratt, Julie
APPLICANT: Reichard, Raymond
APPLICANT: Rosenberg, Martin
APPLICANT: Ward, Judith
TITLE OF INVENTION: Novel Compounds
NUMBER OF SEQUENCES: 1166
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette

```

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/02318
FILING DATE: 19-FEB-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/011,888
FILING DATE: 20-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Glimml, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: GM50006
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 946:
SEQUENCE CHARACTERISTICS:
LENGTH: 1269 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
PCT-US97-02318-946

```

```

Query Match          98.5%;  Score 1250;  DB 1;  Length 1269;
Best Local Similarity 100.0%;  Pred. No. 4.7e-269;
Matches 1269;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

```

QY	1	AAACACCACATCCCAATATGAGAGGAAATCCAAATCAATACCCGAGAACCCAAATCCAAAGTAA	80
Db	1	AAACCCCAATCCCAATATGGGAGGAAATCCCAATCAATACCCGAGAACCCAAATCCAAAGTAA	60
QY	61	TTAAGTCCAAAGTTTGGACAATTACCAAAATATGATTCGATGGGTCAAATGNCANCG	120
Db	61	TTAAGTCCAAAGTTTGGACAATTACCAAAATATGATTCGATGGGTCAAATGNCANCG	120
QY	121	GTGTAAATAACTACGAAATGNTGTGMAAATGATAGANCAAGTTGCCATACAGTNT	180
Db	121	GTGTAAATAACTACGAAATGNTGTGMAAATGATAGANCAAGTTGCCATACAGTNT	180
QY	181	CTCAATTATATGATGCTTTAGAAATCGAATGACACACACAGCCAGTTTACAAATCAATAT	240
Db	181	CTCAATTATATGATGCTTTAGAAATCGAATGACACACACAGCCAGTTTACAAATCAATAT	240
QY	241	TTGTAACTACAAAGATATATAAGAAAGCCGCTATAGACAGAAATGGAAGGTTTGATTTTA	300
Db	241	TTGTAACTACAAAGATATATAAGAAAGCCGCTATAGACAGAAATGGAAGGTTTGATTTTA	300
QY	301	ATGCTCTTAGTAAGAATCATATCAATGAGATGCTATAGTACTACGATTATATTAATT	360
Db	301	ATGCTCTTAGTAAGAATCATATCAATGAGATGCTATAGTACTACGATTATATTAATT	360
QY	361	AAAACCGCTAATTAATTTGTTTTTAGAAAACATATAGTATCAATTTTAAATGTAGTTGACA	420
Db	361	AAAACCGCTAATTAATTTGTTTTTAGAAAACATATAGTATCAATTTTAAATGTAGTTGACA	420
QY	421	TACTAGTACTCAAAATATCTATATACAAATTTCTATATATTTCTTTCGGGGCAGGGTGA	480
Db	421	TACTAGTACTCAAAATATCTATATACAAATTTCTATATATTTCTTTCGGGGCAGGGTGA	480
QY	481	ATTGCCAACGGCAGTAAATAAAGCCGCGACCTGCTAATATGTTCATATTTAGTGGCTG	540
Db	481	ATTGCCAACGGCAGTAAATAAAGCCGCGACCTGCTAATATGTTCATATTTAGTGGCTG	540
QY	541	ATCTAGTACATTTCTAGAGCCGACAGTTAAAGTCTGGATGGGAAAGAATGTTAAATTAT	600
Db	541	ATCTAGTACATTTCTAGAGCCGACAGTTAAAGTCTGGATGGGAAAGAATGTTAAATTAT	600
QY	601	CGACAAAGATATGTAGCGTAATTTGTAAAAATGTATACAAATAGGCTTATTTACGATAA	660

```
|||||
Db 601 CGCAAGATTAATGTGCGATTTGTAATAATGTACAAATAGCGTTATTACATTA 660
QY 661 ATTTTCTCCTTGCATCTTAATTCATGATGAGAGATTTTGTATTAGAGTATCA 720
Db 661 ATTTTCTCCTTGCATCTTAATTCATGATGAGATTTTGTATTAGAGTATCA 720
QY 721 TTGTAGTCAATTAATGATTAATGCGATTCACCTCCAAATAGTGTCAAGGTCAACAG 780
Db 721 TTGTAGTCAATTAATGATTAATGCGATTCACCTCCAAATAGTGTCAACAG 780
QY 781 TGTATTCACCGCTTGGCGCTTGTAGTAAATGAAGTAGAGATTGTGGTATGGTGC 840
Db 781 TGTATTCACCGCTTGGCGCTTGTAGTAAATGAAGTAGAGATTGTGGTATGGTGC 840
QY 841 ACACCTTGAGAAAAAGGTGACACGATCGGAGGTTCAAGCACTTGATATGGCACACANAA 900
Db 841 ACACCTTGAGAAAAAGGTGACACGATCGGAGGTTCAAGCACTTGATATGGCACACANAA 900
QY 901 TGGTGAAGGTGGCAGATTTATTTAGTTAGTAGAGCATGTAGTCATTTGGTCAACACC 960
Db 901 TGGTGAAGGTGGCAGATTTATTTAGTTAGTAGAGCATGTAGTCATTTGGTCAACACC 960
QY 961 ACCCTGTGTTAACAAATTAATGATTTGTAAGTAGCANAAAGTAGTATTACNCAACANAA 1020
Db 961 ACCCTGTGTTAACAAATTAATGATTTGTAAGTAGCANAAAGTAGTATTACNCAACANAA 1020
QY 1021 ACAATTCCTGTAGACACACATGGGTATGACAGCTTACGGGGCTCCACGGTATTAGAGG 1080
Db 1021 ACAATTCCTGTAGACACACATGGGTATGACAGCTTACGGGGCTCCACGGTATTAGAGG 1080
QY 1081 TTGAATTTGGTTGGATGATGAAGGGCATGACAAATTAATACCAAGACTTTTAAAGCA 1140
Db 1081 TTGAATTTGGTTGGATGATGAAGGGCATGACAAATTAATACCAAGACTTTTAAAGCA 1140
QY 1141 AAAAGCAAGCAACTTGCACAAATTTACAGTGAAGTNTCTTGAAGTTTATGATGGTA 1200
Db 1141 AAAAGCAAGCAACTTGCACAAATTTACAGTGAAGTNTCTTGAAGTTTATGATGGTA 1200
QY 1201 AACAAAGCAATTAATGATGAAGACAAAGTCAATGATTAACAAAGAGTTAAACAGA 1260
Db 1201 AACAAAGCAATTAATGATGAAGACAAAGTCAATGATTAACAAAGAGTTAAACAGA 1260
QY 1261 TGTCTATAG 1269
Db 1261 TGTCTATAG 1269

RESULT 2
PCT-US98-24857-3
; Sequence 3, Application PC/TUS9824857A
; GENERAL INFORMATION:
; APPLICANT: Smithline Beecham Corporation
; TITLE OF INVENTION: Novel ribd
; FILE REFERENCE: P50444-9
; CURRENT APPLICATION NUMBER: PCT/US98/24857A
; EARLIER APPLICATION NUMBER: 08/978,456
; EARLIER FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Fastseq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1269
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1269)
; OTHER INFORMATION: n - A,T,C or G
PCT-US98-24857-3
```

Query Match

98.5%; Score 1250; DB 1; Length 1269;

```
Best Local Similarity 100.0%; Pred. No. 4,76-269;
Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
|||||
QY 1 AANCAACCAATCCNATTGGGAGGNAATCCAAATCAATNCCCGGANNCCCATCCAGTTAA 60
Db 1 AANCAACCAATCCNATTGGGAGGNAATCCAAATCAATNCCCGGANNCCCATCCAGTTAA 60
QY 61 TTAAGTCCAGGTTTGGAGAACATTCACCAATATGATTCGATGAGGTCGAATGNCANAG 120
Db 61 TTAAGTCCAGGTTTGGAGAACATTCACCAATATGATTCGATGAGGTCGAATGNCANAG 120
QY 121 GTGTTAATAAATACGAATGNTGTGNAATATGATAGTACANCAAGTTGGCATACAGTT 180
Db 121 GTGTTAATAAATACGAATGNTGTGNAATATGATAGTACANCAAGTTGGCATACAGTT 180
QY 181 CTCNATTATATGATGCTTTAGAAATGCAATGACACACAGCGCAGTTCAATCAATAT 240
Db 181 CTCNATTATATGATGCTTTAGAAATGCAATGACACACAGCGCAGTTCAATCAATAT 240
QY 241 TTGTATAGATTAATTAAGAGAACGCTATATGAGAGCAATTAAGTTGATTTTA 300
Db 241 TTGTATAGATTAATTAAGAGAACGCTATATGAGAGCAATTAAGTTGATTTTA 300
QY 301 ATGCTGTAGTAGAATCATATCATGATGAGATGCCATATAGTACTGAGATTATTAATTA 360
Db 301 ATGCTGTAGTAGAATCATATCATGATGAGATGCCATATAGTACTGAGATTATTAATTA 360
QY 361 AAAACCGTATTAATTTGTTTTTAAAGCAATATGATCAATTTTAAATGATGTTGACA 420
Db 361 AAAACCGTATTAATTTGTTTTTAAAGCAATATGATCAATTTTAAATGATGTTGACA 420
QY 421 TACTAGTCTGCTAATAATCTATTAACAAATTTATATATCTTCCGGGAGGAGTGA 480
Db 421 TACTAGTCTGCTAATAATCTATTAACAAATTTATATATCTTCCGGGAGGAGTGA 480
QY 481 ATTCCCAACCGGACAGTAAATTAAGCCTGCGACCTGCTAATATGTTCAATTTAGTGGCTG 540
Db 481 ATTCCCAACCGGACAGTAAATTAAGCCTGCGACCTGCTAATATGTTCAATTTAGTGGCTG 540
QY 541 ATCTAGTGAATTTTAGACCCGACACTTAAAGTCTGATGAGAGAAAGATTTAATTA 600
Db 541 ATCTAGTGAATTTTAGACCCGACACTTAAAGTCTGATGAGAGAAAGATTTAATTA 600
QY 601 CGCAAGATTAATGATGAGCTATTTGTAATAATGTGACAAATAGGCTTTTAACGATTA 660
Db 601 CGCAAGATTAATGATGAGCTATTTGTAATAATGTGACAAATAGGCTTTTAACGATTA 660
QY 661 ATTTTCTCCTTGCATCTTAATTCATGATGAGAGATTTTGTATTAGAGGTGATCA 720
Db 661 ATTTTCTCCTTGCATCTTAATTCATGATGAGAGATTTTGTATTAGAGGTGATCA 720
QY 721 TTGTAGTCAATTAATGATTAATGCGATTCACCTCCAAATAGTGTCAAGGTCAACAG 780
Db 721 TTGTAGTCAATTAATGATTAATGCGATTCACCTCCAAATAGTGTCAAGGTCAACAG 780
QY 781 TGTATTCACCGCTTGGCGCTTGTAGTAAATGAAGTAGAGATTGTGGTATGGTGC 840
Db 781 TGTATTCACCGCTTGGCGCTTGTAGTAAATGAAGTAGAGATTGTGGTATGGTGC 840
QY 841 ACACCTTGAGAAAAAGGTGACACGATCGGAGGTTCAAGCACTTGATATGGCACACANAA 900
Db 841 ACACCTTGAGAAAAAGGTGACACGATCGGAGGTTCAAGCACTTGATATGGCACACANAA 900
QY 901 TGGTGAAGGTGGCAGATTTATTTAGTTAGTAGAGCATGTAGTCATTTGGTCAACACC 960
Db 901 TGGTGAAGGTGGCAGATTTATTTAGTTAGTAGAGCATGTAGTCATTTGGTCAACACC 960
QY 961 ACCCTGTGTTAACAAATTAATGATTTGTAAGTAGCANAAAGTAGTATTACNCAACANAA 1020
Db 961 ACCCTGTGTTAACAAATTAATGATTTGTAAGTAGCANAAAGTAGTATTACNCAACANAA 1020
QY 1021 ACAATTCCTGTAGACACACATGGGTATGAGAGCTTACGGGGCTCCACGGTATTAGAGG 1080
Db 1021 ACAATTCCTGTAGACACACATGGGTATGAGAGCTTACGGGGCTCCACGGTATTAGAGG 1080
```

Db 1021 acaatccgttagacacacatcgtgtagagcgttacccgggctccacgctatttgagg 1080
QY 1081 TTGAATTCGCTTGATGATGAACGGGCAATCAATTAACCAAGACTTTTTRAAACA 1140
Db 1081 ttgaatttcgcttgatgtagtgaacgggcatcacaattaccgaagactttttaaaga 1140
QY 1141 AAAGCAAGCACTTGGCACAATAATACAGTGAAGTCTTGAAGTTAGATGGGTA 1200
Db 1141 aaagcaagaacttcgcacaaaattcagtgaaagtcttgaagttcagatggta 1200
QY 1201 AACAAAGCAATTGATTAATGACAAAGTCAATGATTAACAAAGAGTTAAACAAGA 1260
Db 1201 aacaagaagtaataatgacaaagccaatgatacttaacaaagagttaaacaaga 1260
QY 1261 TGTCTATAG 1269
Db 1261 tgcctatag 1269

RESULT 3
US-08-903-470-946
; Sequence 946, Application US/08903470
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Burnham, Martin
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Pratt, Julie
; APPLICANT: Reichard, Raymond
; APPLICANT: Rosenberg, Martin
; APPLICANT: Ward, Judith
; TITLE OF INVENTION: Novel Compounds
; NUMBER OF SEQUENCES: 1166
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/903,470
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/011,888
; FILING DATE: 20-FEB-1996
; APPLICATION NUMBER: PCT/US97/02318
; FILING DATE: 19-FEB-1997
; APPLICATION NUMBER: PCT/US97/02547
; FILING DATE: 19-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimm, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: GMS0006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 946:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1269 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; US-08-903-470-946

Query Match 98.5%; Score 1250; DB 26; Length 1269;
Best Local Similarity 100.0%; Pred. No. 4,7e-265;
Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AANACCAATCCNATTTGGAGNATCCAAATCAATNCCCGGANNCCCAATCCAAAGTTAA 60
Db 1 AANACCAATCCNATTTGGAGNATCCAAATCAATNCCCGGANNCCCAATCCAAAGTTAA 60
QY 61 TTAAGTCAAGGTTTGGACATTTACCAATATGATTTCCGATAGGTCAATGNCANCG 120
Db 61 TTAAGTCAAGGTTTGGACATTTACCAATATGATTTCCGATAGGTCAATGNCANCG 120
QY 121 GTGTTAATAACTAGCAATAGNTGTGNAATGATAGAGNCAAGTTGGCCATACAGTNT 180
Db 121 GTGTTAATAACTAGCAATAGNTGTGNAATGATAGAGNCAAGTTGGCCATACAGTNT 180
QY 181 CTCNATTATATGATGCTTTAGAAATCGAATGAGCAACACAGCCAGTTACCAATCAATAT 240
Db 181 CTCNATTATATGATGCTTTAGAAATCGAATGAGCAACACAGCCAGTTACCAATCAATAT 240
QY 241 TTGTAAGTACAGATTAATAAGAGACGCTCTATAGAGCAATGAGGTTGATTTTA 300
Db 241 TTGTAAGTACAGATTAATAAGAGACGCTCTATAGAGCAATGAGGTTGATTTTA 300
QY 301 ATGCTCTGTAGTAGAATCATCAATGAGATGCGCTATAGAGCAATGAGGTTGATTTTA 360
Db 301 ATGCTCTGTAGTAGAATCATCAATGAGATGCGCTATAGAGCAATGAGGTTGATTTTA 360
QY 361 AAAACGCTCAATTAATGTTTTTTTGAAGAAACATATAGTATCATTTAAATGAGTTGACA 420
Db 361 AAAACGCTCAATTAATGTTTTTTTGAAGAAACATATAGTATCATTTAAATGAGTTGACA 420
QY 421 TACTAGTACTCAATATATCTATACAAATTCATATATATTTCTTTGGGGCAGGCTGAA 480
Db 421 TACTAGTACTCAATATATCTATACAAATTCATATATATTTCTTTGGGGCAGGCTGAA 480
QY 481 ATTCCCAACGGCGAGTAATAAAGCCTGCGACCTGCTAATATGTTTCAATTTAGTGCGTG 540
Db 481 ATTCCCAACGGCGAGTAATAAAGCCTGCGACCTGCTAATATGTTTCAATTTAGTGCGTG 540
QY 541 ATCTAGTGAATCTAGAGCCGACAGTTAAAGTCTGATGGAGAAAGAAATGTTAATAT 600
Db 541 ATCTAGTGAATCTAGAGCCGACAGTTAAAGTCTGATGGAGAAAGAAATGTTAATAT 600
QY 601 CGACAAAGATATATGATCGGATTTGTAAGAAATGTTCAAAATAGGCTTATTAACGATA 660
Db 601 CGACAAAGATATATGATCGGATTTGTAAGAAATGTTCAAAATAGGCTTATTAACGATA 660
QY 661 ATTTTCTCCTTGCATCTTAATTCATGATGATGAGATTTTGTGTTATAGAGTGATCA 720
Db 661 ATTTTCTCCTTGCATCTTAATTCATGATGATGAGATTTTGTGTTATAGAGTGATCA 720
QY 721 TTTGAGTCAATTAATGATTAAGATTAAGATTAACCTCCAAATATGTTACAGGAGG 780
Db 721 TTTGAGTCAATTAATGATTAAGATTAAGATTAACCTCCAAATATGTTACAGGAGG 780
QY 781 TGTTAATTCACCCGTTGGCGCTGTGTAGTTAATGAAGTGAAGATTTGTGTTGCTG 840
Db 781 TGTTAATTCACCCGTTGGCGCTGTGTAGTTAATGAAGTGAAGATTTGTGTTGCTG 840
QY 841 ACACCTGAGAAAGGTGACAAAGCATGCGGAGGTTCAAGCACTGATATGCGACAAANAA 900
Db 841 ACACCTGAGAAAGGTGACAAAGCATGCGGAGGTTCAAGCACTGATATGCGACAAANAA 900
QY 901 TGTGAGAGTGCAGCATTAATTAATTAAGTGAAGCCATGAGTCAATTTTGGTTCAACACC 960
Db 901 TGTGAGAGTGCAGCATTAATTAATTAAGTGAAGCCATGAGTCAATTTTGGTTCAACACC 960
QY 961 ACCCTGTGTAAACAAATTAATGATGTTAGATAGCAANAAAGTATTAACCAACANAAG 1020
Db 961 ACCCTGTGTAAACAAATTAATGATGTTAGATAGCAANAAAGTATTAACCAACANAAG 1020


```

Db 1141 AAAAGCAAGCACTTCCCAAAATTTACAGTGAAGTCTTGAAGATTAGATGGGTA 1200
|||||
QY 1201 AACAAAGCAATTCATTAATGAGACAAAGTCATGATTACTTAACAAAGAGTTAAACAAGA 1260
|||||
Db 1201 AACAAAGCAATTCATTAATGAGACAAAGTCATGATTACTTAACAAAGAGTTAAACAAGA 1260
|||||
QY 1261 TGCTATAG 1269
|||||
Db 1261 TGCTATAG 1269

RESULT 5
US-09-611-529-7446/c
: Sequence 7446, Application US/09611529
: GENERAL INFORMATION:
: APPLICANT: George H. Shimer, Jr.
: APPLICANT: George H. Miller
: APPLICANT: Roberta S. Hare
: APPLICANT: Karen J. Shaw
: TITLE OF INVENTION: Staphylococcus aureus Related Compositions and Methods
: FILE REFERENCE: 1034/IC963U1
: CURRENT APPLICATION NUMBER: US/09/611,529
: CURRENT FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/417,811
: PRIOR FILING DATE: 1999-10-14
: PRIOR APPLICATION NUMBER: US 09/353,718
: PRIOR FILING DATE: 1999-07-14
: PRIOR APPLICATION NUMBER: US 09/266,557
: PRIOR FILING DATE: 1999-03-11
: PRIOR APPLICATION NUMBER: US 09/266,556
: PRIOR FILING DATE: 1999-03-11
: PRIOR APPLICATION NUMBER: US 09/266,555
: PRIOR FILING DATE: 1999-03-11
: PRIOR APPLICATION NUMBER: US 09/266,542
: PRIOR FILING DATE: 1999-03-11
: PRIOR APPLICATION NUMBER: US 09/266,541
: PRIOR FILING DATE: 1999-03-11
: PRIOR APPLICATION NUMBER: US 09/037,934
: PRIOR FILING DATE: 1998-03-10
: PRIOR APPLICATION NUMBER: US 09/036,720
: PRIOR FILING DATE: 1998-03-06
: PRIOR APPLICATION NUMBER: US 09/036,338
: PRIOR FILING DATE: 1998-03-06
: PRIOR APPLICATION NUMBER: US 09/036,334
: PRIOR FILING DATE: 1998-03-06
: PRIOR APPLICATION NUMBER: US 09/036,221
: PRIOR FILING DATE: 1998-03-06
: PRIOR APPLICATION NUMBER: US 09/036,137
: PRIOR FILING DATE: 1998-03-06
: PRIOR APPLICATION NUMBER: US 09/036,082
: PRIOR FILING DATE: 1998-03-06
: PRIOR APPLICATION NUMBER: US 09/036,081
: PRIOR FILING DATE: 1998-03-06
: PRIOR APPLICATION NUMBER: US 09/036,079
: PRIOR FILING DATE: 1998-03-06
: PRIOR APPLICATION NUMBER: US 09/035,913
: PRIOR FILING DATE: 1998-03-06
: PRIOR APPLICATION NUMBER: US 09/035,744
: PRIOR FILING DATE: 1998-03-06
: PRIOR APPLICATION NUMBER: US 08/827,356
: PRIOR FILING DATE: 1997-04-01
: PRIOR APPLICATION NUMBER: US 08/831,156
: PRIOR FILING DATE: 1997-04-01
: PRIOR APPLICATION NUMBER: US 60/014,477
: PRIOR FILING DATE: 1996-04-01
: PRIOR APPLICATION NUMBER: US 60/016,743
: PRIOR FILING DATE: 1996-05-02
: PRIOR APPLICATION NUMBER: US 60/020,016
: PRIOR FILING DATE: 1996-06-14
: NUMBER OF SEQ ID NOS: 7451
: SEQ ID NO 7446
: LENGTH: 74105

```

[illegible]

```

OY 1028 CGTTAGACACAAATGGGTGATGTAGAGCTTACGGGGCTCCACGGTATTTGAGGGTTGAATT 1087
      |||||||
Db 23953 CGTTAGACACAAAT-GGTGATGAGAGCTTACGGGGCTCACGGTATTGAGGTTGAATGCGTT 2389
      |||||||
OY 1088 GCGTGTGATGATGAACGGGCGATACACATTAATACCAAGACCTTTTAAACCAAAAAGCA 1147
      |||||||
Db 23894 -----GATGATGACGGCGCATACCAATTTATACCA---AGACTTTTAAACCAAAAAGCA 2384
      |||||||
OY 1148 AAGCAACTGCCACAAATTAATGAGTGAAGATNTCTTGAAGTTTATAGATGGGTAACAAAG 1207
      |||||||
Db 23843 AAGCAACT--GCCCAAAATTAACGATGAAGATATC-TCGCAAGTTTATGATGATGAACCAACG 2378
      |||||||
OY 1208 CGAATTGATTAATGACAAATCAATGAGATTACTACAAAGAGCTTAACACAGATGTCTAT 1267
      |||||||
Db 23786 ---AATGATTAATGACAAAGCTCAATGATTACTTAACAAAGAGCTTAACACAGATGTCTAT 2373
      |||||||
OY 1268 A 1268
Db 23729 A 23729

RESULT 6
US-08-831-156A-107/c
Sequence 107, Application US/08831156A
GENERAL INFORMATION:
  APPLICANT: George H. Shimer, Jr.
  APPLICANT: George H. Miller
  APPLICANT: Roberta S. Hare
  APPLICANT: Karen J. Shaw
  TITLE OF INVENTION: NUCLEOTIDE SEQUENCE OF A
  TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS
  TITLE OF INVENTION: GENOME AND RELATED
  NUMBER OF SEQUENCES: 112
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: Schering-Plough Corporation
  STREET: 2000 Galloping Hill Road
  CITY: Kenilworth
  STATE: New Jersey
  COUNTRY: USA
  ZIP: 07033-0530
  COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn
  CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/831,156A
  FILING DATE: 01-APR-1997
  PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 60/014,477
  FILING DATE: 01-APR-1996
  APPLICATION NUMBER: 60/016,743
  FILING DATE: 02-MAY-1996
  APPLICATION NUMBER: 60/020,016
  FILING DATE: 14-JUN-1996
  INFORMATION FOR SEQ. ID NO: 107:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 8517 base pairs
  TYPE: nucleic acid
  STRANDEDNESS: double
  TOPOLOGY: circular
  MOLECULE TYPE: DNA (genomic)
  HYPOTHEICAL: NO
  ANTI-SENSE: NO
  ORIGINAL SOURCE:
  ORGANISM: Staphylococcus aureus
  FEATURE:
  NAME/KEY: misc.feature
  LOCATION: 1...85161
US-08-831-156A-107

```

Query Match	76.1%	Score 965.6	DB 23	Length 85137
Best Local Similarity	93.5%	Pred. No. 5.2e-203		
Matches 1123	Conservative	0	Mismatches 57	Indels 21
				Gaps 11
QY	68	CAGGCTTTTGGAACTTTCACCAATATGATCCGATGAGGCTCAATGNCANCGTGTAA	127	
DB	50771	CCAAGCTTTGGAACTTAAACAATATGATTTGGAT-AGGTCAATGACAAACGGCTTTAA	50713	
QY	128	TAACTACGAATGNTGTGMAATGATAGTACAMCAAGTTGGCCATACAGTCTCMAAT	187	
DB	50712	TAACTACGAATGNGTGAATATGATATAGAACAAAGTTGGCCATACAGTCTCMAAT	50653	
QY	188	ATATGATGCTTTAGATCGAATGAGCAACAACAGCGCACTTCAATCATATTTGTAC	247	
DB	50652	ATATGATGCTTTAGATCGAATGAGCAACAACAGCGCACTTCAATCATATTTGTAC	50593	
QY	248	TGAAGATATTAAGAAACGCTCTATAGACGAATTAAGGTTTGATTTAAATGTCTG	307	
DB	50592	TGAAGATATTAAGAAACGCTCTATAGACGAATTAAGGTTTGATTTAAATGTCTG	50533	
QY	308	TTAGTAAATCATATCAATGAGATGCTTATAGTACTCAGATTATATTAATTAACCG	367	
DB	50532	TTAGTAAATCATATCAATGAGATGCTTATAGTACTCAGATTATATTAATTAACCG	50473	
QY	368	TCATTAATTTGTTTTTTAGAAAACATATATGATTCATTTTAATGTATGTCACATCTAG	427	
DB	50472	TCATTAATTTGTTTTTTAGAAAACATATATGATTCATTTTAATGTATGTCACATCTAG	50413	
QY	428	TACTCAATATATCTTATACCAATTTTCATATATATCTTTTCGGGGCAGGGTGAATCCCA	487	
DB	50412	TACTCAATATATCTTATACCAATTTTCATATATATCTTTTCGGGGCAGGGTGAATCCCA	50353	
QY	488	ACCGGCAGTAAATAAAGCCCTGCGACCTGCTAATATGTTTTCAATATAGTGGCTGATCTAG	547	
DB	50352	ACCGGCAGTAAATAAAGCCCTGCGACCTGCTAATATGTTTTCAATATAGTGGCTGATCTAG	50293	
QY	548	GAGATTTTACAGCCGCACAGTTAAAGTCTGATGGGAAAGAAATGTAATATTCGACAA	607	
DB	50292	GAGATTTTACAGCCGCACAGTTAAAGTCTGATGGGAAAGAAATGTAATATTCGACAA	50233	
QY	608	GATTAATGTAGCTATTTGTAAAAATGTGACAAATAGGCTTATTAACGATAAATTTTC	667	
DB	50232	GATTAATGTAGCTATTTGTAAAAATGTGACAAATAGGCTTATTAACGATAAATTTTC	50173	
QY	668	TGCTTTGCATCTTAATTCATGATGTGAGGATTTTGTATATAGAGGTGATCTTTGAGT	727	
DB	50172	TGCTTTGCATCTTAATTCATGATGTGAGGATTTTGTATATAGAGGTGATCTTTGAGT	50115	
QY	728	CAATTTATGATTTATGCGATTCACCTCCCAATATGATACAAAGTCCANACAGGTGTAAT	787	
DB	50114	CAATTTATGATTTATGCGATTCACCTCCCAATATGATACAAAGTCCANACAGGTGTAAT	50055	
QY	788	CCACCCGTTGGCGCTGTGTACTTAAATGAAGTAGGATGTTGCTATGCTGCACCTTG	847	
DB	50054	CCACCCGTTGGCGCTGTGTACTTAAATGAAGTAGGATGTTGCTATGCTGCACCTTG	49995	
QY	848	AGAAAGGTGACAAACATGCGGAGGTTCACGACTTGATATGGCAACAAANATGCTGAA	907	
DB	49994	AGAAAGGTGACAAACATGCGGAGGTTCACGACTTGATATGGCAACAAANATGCTGAA	49935	
QY	908	GGTGGCAGATTTATATTACGTTAGAGCCATGTAGTCAATTTTGTTCACACCCCTGT	967	
DB	49934	GGTGGCAGATTTATATTACGTTAGAGCCATGTAGTCAATTTTGTTCACACCCCTGT	49875	
QY	968	GTTAACAAATATATGATTTGTAAGTATAGCAANAAGTATTTACNCAACAAAGACATTC	1027	
DB	49874	GTTAACAAATATATGATTTGTAAGTATAGCAANAAGTATTTACNCAACAAAGACATTC	49815	
QY	1028	CGTTGACACACATGGGTGATGAGACGTTACGGGGCTCCACGCTTTTACGAGGTTAAAT	1087	
DB	49816	CGTTGACACACAT-GGGTGATGAGACGTTACGGGGCTCCACGCTATTTGAGGTTGAATCGTT	49758	
QY	1088	GGGTTGGATGATGAACGGGCATCACAATTTATCCAAAGCTTTTTTTTAAAGCAAAAAGA	1147	


```
RESULT
US-08-956-171-142/C
; Sequence 142, Application US/08956171
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 142:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7588 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
US-08-956-171-142

Query Match          75.5%  Score 957.6:  DB 30:  Length 7588;
Best Local Similarity 93.1%  Pred. No. 1.6e-203;
Matches 1118:  Conservative 0; Mismatches 62; Indels 21; Gaps 11;
```

```

Db 6153 TACTCAATAATCTATACAAATTTTCATATATATCTTTGGGGGAGGGTGAATTCCTCA 6094
QY 488 ACCGCGAGTAATAAAGCGCGACCTGCTAATATGTTTCATATATGTCGATCTAGT 547
Db 6093 ACCGCGAGTAATAAAGCGCGACCTGCTAATATGTTTCATATATGTCGATCTAGT 6034
QY 548 GAGATCTGAGCGGACAGTAAAGTCTGATGGAGGAAGAATGTTAATATCGCAAA 607
Db 6033 GAGATCTGAGCGGACAGTAAAGTCTGATGGAGGAAGAATGTTAATATCGCAAA 5974
QY 608 GATAATGACGCTATTGTTAAAGTGTACAAATAGGCTTATTACATTAATTTTC 667
Db 5973 GATAATGACGCTATTGTTAAAGTGTACAAATAGGCTTATTACATTAATTTTC 5914
QY 668 TCCTTGCATCTTAATTCATGATGAGATTTTGTATTATAGCTGATATCTAGT 727
Db 5913 TCC-TTGCAATCTTAATTCATGATGAGTGAAGA-TTTTGTATTATAGCTGATATCTAGT 5856
QY 728 CAATTATGATATGCGATTCACCTTCCAAATATGTTACAAAGTCCANACAGCTGTTAT 787
Db 5855 CAATTATGATATGCGATTCACCTTCCAAATATGTTACAAAGTCCANACAGCTGTTAT 5796
QY 788 CCACCCGTTGGCGCTGTGTAGTTAATGAAGTAGAGTGTGTTATGTCACACTTG 847
Db 5795 CCACCCGTTGGCGCTGTGTAGTTAATGAAGTAGAGTGTGTTATGTCACACTTG 5736
QY 848 AGAAAGGTGACAGCATCGGAGGTTCAAGCACTTGATATGGCCACAAANATGCTGAA 907
Db 5735 AGAAAGGTGACAGCATCGGAGGTTCAAGCACTTGATATGGCCACAAANATGCTGAA 5676
QY 908 GGTGCGACGATTTATTTACGTTAGAGCATGTTGTTGTTCAACACACCTGCT 967
Db 5675 GGTGCGACGATTTATTTACGTTAGAGCATGTTGTTGTTCAACACACCTGCT 5616
QY 968 GTTAAACAAATATTGATTGTGAAGATGACAAATGATTTACNCACAAAGACAATTC 1027
Db 5615 GTTAAACAAATATTGATTGTGAAGATGACAAATGATTTACNCACAAAGACAATTC 5556
QY 1028 CGTTACACACATGGGTGATGAGACGTTACGGGCTCCACGGTATTGAGGTTGAATTC 1087
Db 5557 CGTTACACACATGGGTGATGAGACGTTACGGGCTCCACGGTATTGAGGTTGAATTC 5499
QY 1088 GCGTTGATGATGAACGGGATCACAATTTATACAAAGACTTTTAAAGCAAAAAGCA 1147
Db 5498 -----GATGATGAACGGGATCACAATTTATACCA---AGACTTTTAAAGCAAAAAGCA 5448
QY 1148 AAGCAACTGGCCACAAATTTACAGTGAAGTCTTTGAAGTTTACATGGGTTAAACAAG 1207
Db 5447 AAGCAACT--GCCACAAATTTACAGTGAAGTATC-TGCAAGTTTACATGGGTTAAACAAG 5391
QY 1208 CGAATTGATTAATGGAACAAAGTCAATGATTTACTAACAAAGAGGTTAAACAAGATGCTAT 1267
Db 5390 ---AATGATTAATGGAACAAAGTCAATGATTTACTAACAAAGAGGTTAAACAAGATGCTAT 5334
QY 1268 A 1268
Db 5333 A 5333

RESULT
US-08-956-171B-142/C
; Sequence 142, Application US/08956171B
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gail H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
;
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
```

STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956.171B
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
Prior APPLICATION DATA:
APPLICATION NUMBER: 60/009.861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Hoover, Kenley K.
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 610-5790
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 142:
SEQUENCE CHARACTERISTICS:
LENGTH: 7588 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 142:
US-08-956-171B-142

Query Match 75.5%; Score 957.6; DB 30; Length 7588;
Best Local Similarity 93.1%; Pred. No. 1.6e-203;
Matches 1118; Conservative 0; Mismatches 62; Indels 21; Gaps 11;

QY 68 CAAAGTTTGGACATTTACCAATATGATTCGGATGAGTCAATGNCANCGGTGTTAA 127
DB 6512 CCAAGGTTGGACATTTACCAATATGATTCGGAT-AGGTCAATGACAAACGGGTGTTAA 6454
QY 128 TAACTAGCAATGNTGNAATGATAGTAGNCAAGTGGCATACAGTTCCTCNAAT 187
DB 6453 TAACTAGCAATGNTGNAATGATAGTAGNCAAGTGGCATACAGTTCCTCNAAT 6394
QY 188 ATATGATCCTTAGAATGCAATGAGCAACACAGCGAGTTACATCAATATTTGTAAC 247
DB 6393 ATATGATCCTTAGAATGCAATGAGCAACACAGCGAGTTACATCAATATTTGTAAC 6334
QY 248 TAGAAGATTAATAGAGACCGCTATAGAGACGATTTGAAGTTGATTTTATGCTG 307
DB 6333 TAGAAGATTAATAGAGACCGCTATAGAGACGATTTGAAGTTGATTTTATGCTG 6274
QY 308 TTAGTAAGATCAATTCATGATGAGTGGCTATGCTAGTCAATTAATTAATAAACG 367
DB 6273 TTAGTAAGATCAATTCATGATGAGTGGCTATGCTAGTCAATTAATTAATAAACG 6214
QY 368 TCATTAAATGTTTTTTCGAAAACATATAGTATCAATTTAATAGTGTGACATCTACG 427
DB 6213 TCATTAAATGTTTTTTCGAAAACATATAGTATCAATTTAATAGTGTGACATCTACG 6154
QY 428 TACTCAATTAATCTATTAACATTTTCAATATATATTTCTTGGGGGAGGTGAATTTCCA 487
DB 6153 TACTCAATTAATCTATTAACATTTTCAATATATATTTCTTGGGGGAGGTGAATTTCCA 6094
QY 488 ACCGGCAGTAATAAGACCTGGAGCTGCTATATGTTTCAATATAGTGGCTGATCTAGT 547
DB 6093 ACCGGCAGTAATAAGACCTGGAGCTGCTATATGTTTCAATATAGTGGCTGATCTAGT 6034
QY 548 GAGATTCTAGAGCCGACGATTAAAGTCTGATGGAGAAAGATGTTAATATCGACAAA 607

DB 6033 GAGATTCTAGAGCCGACGATTAAAGTCTGATGGAGAAAGATGTTAATATCGACAAA 5974
QY 608 GATAATGAGCGTATTTGCTAAATATGTTACAAATAGCGTTATTTACGATTAATTTTC 667
DB 5973 GATAATGAGCGTATTTGCTAAATATGTTACAAATAGCGTTATTTACGATTAATTTTC 5914
QY 668 TCCTTTCATCTTAATTCATGATGAGGATTTTGTGTTATAGAGGTGATCATTTGAGT 727
DB 5913 TCC-TTGATCTTAATTCATGATGAGGATTTTGTGTTATAGAGGTGATCATTTGAGT 5856
QY 728 CAATTTATGATTTATGCGATTCAACTTCCAAATATGTTACAAAGTTCANACAGTGTTAAT 787
DB 5855 CAATTTATGATTTATGCGATTCAACTTCCAAATATGTTACAAAGTTCANACAGTGTTAAT 5796
QY 788 CCACCGGTGGCGCTGTTGATTAATGAAGGTGATGTTGTTGTTGGTGGCACACTTG 847
DB 5795 CCACCGGTGGCGCTGTTGATTAATGAAGGTGATGTTGTTGTTGGTGGCACACTTG 5736
QY 848 AGAAAGGTGACACATGCGAGGTTCAAGCACTTGTATGCGACACANAAATGCTGAA 907
DB 5735 AGAAAGGTGACACATGCGAGGTTCAAGCACTTGTATGCGACACANAAATGCTGAA 5676
QY 908 GGTCCGAGATTTATATAGTATGACCATGATGATTTGTTGTTGTTGTTGTTGTTGTTGTT 967
DB 5675 GGTCCGAGATTTATATAGTATGACCATGATGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 5616
QY 968 GTTAAACAAATTTATGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1027
DB 5615 GTTAAACAAATTTATGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 5556
QY 1028 CGTTAGACACATGAGTGGTATGAGACGTTACGGGCTCCACGATTTGTTGAGGGTTGAT 1087
DB 5557 CGTTAGACACATGAGTGGTATGAGACGTTACGGGCTCCACGATTTGTTGAGGGTTGAT 5499
QY 1088 GCGTTGATGATGATGAGGCGCATCACAATTTATCCAAAGACTTTTAAAGCAAAAAGCA 1147
DB 5498 GCGTTGATGATGATGAGGCGCATCACAATTTATCCAAAGACTTTTAAAGCAAAAAGCA 5448
QY 1148 AAGCACTTGGCAAAAATTTACAGTGAAGGTTGAAAGTTAGATGGGTAAACAAG 1207
DB 5447 AAGCACTTGGCAAAAATTTACAGTGAAGGTTGAAAGTTAGATGGGTAAACAAG 5381
QY 1208 CGAATTTGATTAAGCAAAAGTCAATGATTTACTTAACAAGAGGTTAAACAAGATGCTAT 1267
DB 5380 CGAATTTGATTAAGCAAAAGTCAATGATTTACTTAACAAGAGGTTAAACAAGATGCTAT 5334
QY 1268 A 1268
DB 5333 A 5333

RESULT 10
US-60-038-697-616
Sequence 616, Application US/60038697
GENERAL INFORMATION:
APPLICANT: Lagace, Robert E.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF STAPHYLOCOCCUS AUREUS
TITLE OF INVENTION: GENOME, FRAGMENTS THEREOF, AND USBS THEREOF
NUMBER OF SEQUENCES: 1027
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/60/038,697
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: CERRONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PO-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 855-0555
TELEFAX: (415) 845-4166
INFORMATION FOR SEO ID NO: 616:
SEQUENCE CHARACTERISTICS:
LENGTH: 7003 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: SAC0616
US-60-038-697-616

Query Match      73.4%: Score 931.6: DB 65: Length 7003:
Best Local Similarity 92.2%: Pred. No. 1e-197:
Matches 1109: Conservative 0: Mismatches 72: Indels 22: Gaps 12:

QY 68 CAAGGTTTGGAAACATTACCAAAATATGATTCGATGAGGTCAATGNCANCGGTGTAA 127
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2138 CCAAGTTTGGAAACATTACCAAAATATGATTCGAT -AGTCAAAATGACAAACGGTGTAA 2196

QY 128 TAAACTCGAAATGNTGTGAAAATGATAGTAGANCAGTTCGGATACAGTNTCTCAAT 187
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2197 TAAACTCGAAATGNTGTGAAAATGATAGTAGAGCAAGTTCGGATACAGTNTCTCAAT 2256

QY 188 ATATGATGCTTTAGATGCGATGAGCAACAACAGCGAGTTACATCAATATTTGTAA 247
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2257 ATATGATGCTTTAGATGCGATGAGCAACAACAGCGAGTTACATCAATATTTGTAA 2316

QY 248 TAGAAGATAATAAAGAGACGCTCTATAGAGACGAAATTTGAAGTTTAAATGTCTG 307
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2317 TAGAAGATAATAAAGAGACGCTCTATAGAGACGAAATTTGAAGTTTAAATGTCTG 2376

QY 308 TTAGTAGAATCATATTAATGAGATGCTATAGTACATCAATTTATTAATTTAAACG 367
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2377 TTAGTAGAATCATATTAATGAGATGCTATAGTACATCAATTTATTAATTTAAACG 2436

QY 368 TCATTAATGTTTTTTTGAAGAAACATATAGTATCATTTTAAATGTAGTGCATCTACG 427
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2437 TCATTAATGTTTTTTTGAAGAAACATATAGTATCATTTTAAATGTAGTGCATCTACG 2496

QY 428 TACTCAAAATATCTATACAAATTTTCATATATTAATTTCTTGGGGCAGGGTGAATTC 487
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2497 TACTCAAAATATCTATACAAATTTTCATATATTAATTTCTTGGGGCAGGGTGAATTC 2556

QY 488 ACCGGCAGTAAATAAACCCCTGCCAGCTGCTATATGTTTTCATATTAATGCTGAT 547
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2557 ACCGGCAGTAAATAAACCCCTGCCAGCTGCTATATGTTTTCATATTAATGCTGAT 2616

QY 548 GAGATTGTAGAGCGAGAGTTAAAGTCTGATGAGAGAAAGATGTAATATGAGACAA 607
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2617 GAGATTGTAGAGCGAGAGTTAAAGTCTGATGAGAGAAAGATGTAATATGAGACAA 2676

QY 608 GATAATGTACGATTTTGAATAATGTGTACAAATAGGCTTATTTAA -CGATAAATTTT 666
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2677 GATAATGTACGATTTTGAATAATGTGTACAAATAGGCTTATTTAA -CGATAAATTTT 2736

QY 667 CTCCTTTGCAT -CTTATTTCAATGATGAGAGATTTTGTTTTATAGAGGATCATATTGA 725
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2737 CCCCTTTGCATCTTATTTTATGATGAGAG -TTTGTGTTATAGAGGATCATATTGA 2795

QY 726 GTCAATTTATGATTAATGCGATTCACTTCCAAATATGATGATCAAGGTCAANACAGGT 785

```

```

DB 2796 GTCAATTTATGATTAATGCGATTTCGATTCACATTCGCAAAATATGATGACAAAGGTCA 2855
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 786 ATCCACCCCTGGCCCTGTTGTATGATTAATGAAAGTATGATGTTGTATGGTGCACACT 845
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2856 ATCCACCCCTGGCCCTGTTGTATGATTAATGAAAGTATGATGTTGTATGGTGCACACT 2915
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 846 TGAGAAAGGTGACAAAGCATGCGAGGTTCAACACTTGTATGATGACAAACAAATGCTG 905
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2916 TGAGAAAGGTGACAAAGCATGCGAGGTTCAACACTTGTATGATGACAAACAAATGCTG 2975
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 906 AAGTGGCAGATTTATATTACGTTAGAGCCATGATGTCATTTTGGTTCAACACACCT 965
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2976 AAGTGGCAGATTTATATTACGTTAGAGCCATGATGTCATTTTGGTTCAACACACCT 3035
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 966 GTGTTAACAAATATTGATTTGTAAGATGACAAAGATGATTTATACCAACAAAGACAT 1025
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3036 GTGTTAACAAATATTGATTTGTAAGATGACAAAGATGATTTATACCAACAAAGACAT 3094
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1026 TCCGTTAGACACATGGGATGATGAGACGTTAGGGGCTCCAGGATTTGAGGGTTGAA 1085
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3095 T-CGTTAGATACACAT -GGTATGATACGTTACGGGATCCGTTATGAGGTTGAATGCG 3152
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1086 TTGCGTTGATGATGAACGGGATCAACAATTAATACCAAAAGCTTTTAAACCAAAAG 1145
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3153 TT-----GATGATGAACGTTCAACAATTAATACCA--AGACTTTTAAAGCAAAAG 3203
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1146 CAAGCACTGTCACAAATTTACAGTGAAGTNTCTTGAAGTTTATGATGGTAAACAA 1205
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3204 CAAGCACTGTCACAAATTTACAGTGAAGTNTCTTGAAGTTTATGATGGTAAACAA 3260
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1206 AGCGAATGTAATGGAACAAAGTCAATGATGATTTACTTACAAAGAGGTTAAACAGATG 1265
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3261 CG---ATGTAATTTGACAAAGTCAATGATGATTTACTTACAAAGAGGTTAAACAGATG 3317
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1266 ATA 1268
   |||
DB 3318 ATA 3320

```

```

RESULT 11
US-60-046-714-618
; Sequence 618, Application US/60046714
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert E.
; APPLICANT: Coley, Neil C.
; APPLICANT: Russo, Frank D.
; APPLICANT: Health, Joe D.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF STAPHYLOCOCCUS AUREUS
; TITLE OF INVENTION: GENOME, FRAGMENTS THEREOF, AND USES THEREOF
; NUMBER OF SEQUENCES: 1050
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/046,714
; FILING DATE: HERewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PM-0001-1P P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 855-0555

```


TELEFAX: (415) 845-4166
: INFORMATION FOR SEQ ID NO: 618:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 7004 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: IMMEDIATE SOURCE:
: CLONE: SAU1C618
US-60-046-714-618

Query Match 73.4%; Score 931.6; DB 67; Length 7004;
Best Local Similarity 92.2%; Pred. No. 1e-197;
Matches 1109; Conservative 0; Mismatches 72; Indels 22; Gaps 12;

Oy 68 CAAGTTTGGAACTTACCAATATGATTCGATGAGTCAATGNCANCGGTGTTAA 127
Db 2138 CCAAGTTTGGAACTTACCAATATGATTCGAT-AGGTCAAAATGACAAAGGTGTTAA 2196
Oy 128 TAACTAGCAATGNTGNNAAATGATAGANCAAGTGGCCATACAGTTCNAT 187
Db 2197 TAACTAGCAATGNTGNNAAATGATAGANCAAGTGGCCATACAGTTCNAT 2256
Oy 188 ATATGATGCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 247
Db 2257 ATATGATGCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 2316
Oy 248 TAGAAGATATTAAGAAGAACGCTATATAGAGAGCAATTAAGTTGATTTATGCTG 307
Db 2317 TAGAAGATATTAAGAAGAACGCTATATAGAGAGCAATTAAGTTGATTTATGCTG 2376
Oy 308 TTAGTAAGAT 367
Db 2377 TTAGTAAGAT 2436
Oy 368 TCATTAATTTGTTTTTGTAGAAAACATATATATATATATATATATATATAT 427
Db 2437 TCATTAATTTGTTTTTGTAGAAAACATATATATATATATATATATATATAT 2496
Oy 428 TACTCAAT 487
Db 2497 TACTCAAT 2556
Oy 488 ACCGCGAGTAATTAAGAGCTGCGAGCTGCTATATATATATATATATATAT 547
Db 2557 ACCGCGAGTAATTAAGAGCTGCGAGCTGCTATATATATATATATATATAT 2616
Oy 548 GAGATTTAGAGCCGACAGTTTAAAGTCTGATGGAGAAAGATTTAATTTCGACAA 607
Db 2617 GAGATTTAGAGCCGACAGTTTAAAGTCTGATGGAGAAAGATTTAATTTCGACAA 2676
Oy 608 GATAATGATGCTATTTGTAATAATGCTGTAATTTTAA-CGATTAATTTTTT 666
Db 2677 GATAATGATGCTATTTGTAATAATGCTGTAATTTTAA-CGATTAATTTTTT 2736
Oy 667 CTCCTTTGAT-CTTAATTCATGATGCTGAGTTTCTTATATAGAGTGCATTTGA 725
Db 2737 CTCCTTTGAT-CTTAATTCATGATGCTGAGTTTCTTATATAGAGTGCATTTGA 2795
Oy 726 GTCATTTATGATTTATGATTTATGATTTATGATTTATGATTTATGATTTAT 785
Db 2796 GTCATTTATGATTTATGATTTATGATTTATGATTTATGATTTATGATTTAT 2855
Oy 786 ATCAACCCGTTGGCGCTGTTAGTTAATGAAGTAGAGTTGTTGTTATGTCACACT 845
Db 2856 ATCAACCCGTTGGCGCTGTTAGTTAATGAAGTAGAGTTGTTGTTATGTCACACT 2915
Oy 846 TGAGAAAAGGTACAAGCATGGGAGGTTCAAGCATTTATATGTCACAAACAAATGCTG 905
Db 2916 TGAGAAAAGGTACAAGCATGGGAGGTTCAAGCATTTATGTCACAAACAAATGCTG 2975

Oy 906 AAGTGCGCAGCATTTATATATACGTTAGAGCCATGTAGTCAATTTGGTTCAACACCCT 965
Db 2976 AAGTGCGCAGCATTTATATATACGTTAGAGCCATGTAGTCAATTTGGTTCAACACCCT 3035
Oy 966 GTGTTAACAAATTTATTTATGTTAAGATAGCANAAGTGTATTCNCNACAAAGACAT 1025
Db 3036 GTGTTAACAAATTTATTTATGTTAAGATAGCANAAGTGTATTCNCNACAAAGACAT 3094
Oy 1026 TCCGTTAGACACATGCGTATGATGATGATGATGATGATGATGATGATGATGAT 1085
Db 3095 T-CGTTATATACACT-GGTATATGATGATGATGATGATGATGATGATGATGATGAT 3152
Oy 1086 TTGCGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1145
Db 3153 TT-----GATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3203
Oy 1146 CAAGCACTTGCCACAAATTTACAGTAAAGTTCCTTGAAGTTTATGATGGTTAAACA 1205
Db 3204 CAAGCACT--GCCACAAATTTACAGTAAAGTATC-TGCAAGTTTATGATGTTAAACAAG 3260
Oy 1206 AGCGAATGATATGAGACAAAGTCAATGATTTACTTAACCAAGAGTTTAAACAAGTGTCT 1265
Db 3261 CG---AATGATATATGACAAAGTCAATGATTTACTTAACCAAGAGTTTAAACAAGTGTCT 3317
Oy 1266 ATA 1268
Db 3318 ATA 3320

RESULT 12
US-09-611-529-2079
: Sequence 2079, Application US/09611529

: GENERAL INFORMATION:

: APPLICANT: George H. Shimer, Jr.

: APPLICANT: George H. Miller

: APPLICANT: Karen J. Shaw

: APPLICANT: Roderica S. Hare

: TITLE OF INVENTION: Staphylococcus aureus Related Compositions and Methods

: FILE REFERENCE: 1034/1C963US1

: CURRENT APPLICATION NUMBER: US/09/611,529

: CURRENT FILING DATE: 2000-06-30

: PRIOR APPLICATION NUMBER: US 09/417,811

: PRIOR FILING DATE: 1999-10-14

: PRIOR APPLICATION NUMBER: US 09/353,718

: PRIOR FILING DATE: 1999-07-14

: PRIOR APPLICATION NUMBER: US 09/266,557

: PRIOR FILING DATE: 1999-03-11

: PRIOR APPLICATION NUMBER: US 09/266,556

: PRIOR FILING DATE: 1999-03-11

: PRIOR APPLICATION NUMBER: US 09/266,555

: PRIOR FILING DATE: 1999-03-11

: PRIOR APPLICATION NUMBER: US 09/266,542

: PRIOR FILING DATE: 1999-03-11

: PRIOR APPLICATION NUMBER: US 09/266,541

: PRIOR FILING DATE: 1999-03-11

: PRIOR APPLICATION NUMBER: US 09/037,934

: PRIOR FILING DATE: 1998-03-10

: PRIOR APPLICATION NUMBER: US 09/036,720

: PRIOR FILING DATE: 1998-03-06

: PRIOR APPLICATION NUMBER: US 09/036,338

: PRIOR FILING DATE: 1998-03-06

: PRIOR APPLICATION NUMBER: US 09/036,334

: PRIOR FILING DATE: 1998-03-06

: PRIOR APPLICATION NUMBER: US 09/036,221

: PRIOR FILING DATE: 1998-03-06

: PRIOR APPLICATION NUMBER: US 09/036,137

: PRIOR FILING DATE: 1998-03-06

: PRIOR APPLICATION NUMBER: US 09/036,082

: PRIOR FILING DATE: 1998-03-06

: PRIOR APPLICATION NUMBER: US 09/036,081

: PRIOR FILING DATE: 1998-03-06

: PRIOR APPLICATION NUMBER: US 09/036,079

: PRIOR FILING DATE: 1998-03-06

;; PRIOR APPLICATION NUMBER: US 09/035,913
;; PRIOR FILING DATE: 1998-03-06
;; PRIOR APPLICATION NUMBER: US 09/035,744
;; PRIOR FILING DATE: 1998-03-06
;; PRIOR APPLICATION NUMBER: US 08/827,356
;; PRIOR FILING DATE: 1997-04-01
;; PRIOR APPLICATION NUMBER: US 08/831,156
;; PRIOR FILING DATE: 1997-04-01
;; PRIOR APPLICATION NUMBER: US 60/014,477
;; PRIOR FILING DATE: 1996-04-01
;; PRIOR APPLICATION NUMBER: US 60/016,743
;; PRIOR FILING DATE: 1996-05-02
;; PRIOR APPLICATION NUMBER: US 60/020,016
;; PRIOR FILING DATE: 1996-06-14
;; NUMBER OF SEQ ID NOS: 7451
;; SEQ ID NO: 2078
;; LENGTH: 1119
;; TYPE: DNA
;; ORGANISM: Staphylococcus aureus
US-09-611-529-2079

Query Match 32.9%; Score 417.2; DB 57; Length 1119;
Best Local Similarity 89.7%; Pred. No. 4,8e-83;
Matches 560; Conservative 0; Mismatches 44; Indels 20; Gaps 10;

OY 645 GCTTTATTACATAAATTTCTCTTGCACTTAATTCATGATGATGATTTTGG 704
DB 1 gcttattcaagataattcttcc-ttcgacattcaatgagtgagga-ttttgg 58
OY 705 TTTATAGAGGTATCATTTGAGTCAATTTATGATTCGATTCACCTCCAAATATGG 764
DB 59 ttatagagtgatcatcttgagtgcaattatgatatcgatcacttgcaaatatg 118
OY 765 TACAAGGTCAACAGGTGTTATTCACCCGTTGGCGCTGTTGTTAATGAAGTAGGA 824
DB 119 tacaaggtcaacaggtgtaattccaccgcttgcgctgttgagtaataagtagga 178
OY 825 TTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 884
DB 179 ttgttggttggttggttggttggttggttggttggttggttggttggttggttg 238
OY 885 ATATGGCACAAANAAATGCTGAAGTGCAGCATTTATACGTAGAGCCATGTATGC 944
DB 239 atatggcacaaacaaatgctgaagtgctgagcatattatattagtgagccatgtatgc 298
OY 945 ATTTGGTTCACACACCCTGTGTTAACAAAATTATGATTGTAAGATAGCANAAAGTAG 1004
DB 299 atttgggttcacacaccctgtgttaacaaaattatgattgtaagtagcnaaagtag 358
OY 1005 TATTACNCAACANAAATTCCTGTATACACACATGGGTATGAGAGTTACGGGGCT 1064
DB 359 ta-taagcaacaagaacaaat-cgttagacacat-9gtgtagaggttaccggtc 415
OY 1065 CCACGGTATTAGAGGTTCATTTGCTTGATGATGAAGGCGATCACAATTATACAAA 1124
DB 416 acggtatctaggttgatcgtt-----gatgataaagggcatcaccatacaca-- 467
OY 1125 GACTTTTTTAAAGCAAAAGCAAAAGCACTTGCCACAAATTACAGTGAAGTNTCTTG 1184
DB 468 -agactttttaagaagaagaagaagaagcaact--gccacaattacagtgaaagtatc-tg 523
OY 1185 AAGATTATATGGTGAAGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1244
DB 524 caagttagatgtagtaacaagcg--aatgataatgacaagtaacatgattactaaca 580
OY 1245 AAGAGGTAAACAGATGCTATTA 1268
DB 581 aagaggttaacaagaatgcttata 604

RESULT 13
US-08-827-356-1376

;; Sequence 1376, Application US/08827356
;; GENERAL INFORMATION:
;; APPLICANT: George H. Shimer, Jr.
;; APPLICANT: George H. Miller
;; APPLICANT: Robert S. Hare
;; APPLICANT: Karen J. Shaw
;; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS RELATED
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS
;; NUMBER OF SEQUENCES: 5574
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Schering-Plough Corporation
;; STREET: 2000 Galloping Hill Road
;; CITY: Kenilworth
;; STATE: New Jersey
;; COUNTRY: USA
;; ZIP: 07033-0530
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/827,356
;; FILING DATE: 01-APR-1997
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 60/014,477
;; FILING DATE: 01-APR-1996
;; APPLICATION NUMBER: 60/016,743
;; FILING DATE: 02-MAY-1996
;; APPLICATION NUMBER: 60/020,016
;; FILING DATE: 14-JUN-1996
;; INFORMATION FOR SEQ ID NO: 1376:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1119 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: circular
;; MOLECULE TYPE: DNA (genomic)
;; HYPOTHEICAL: NO
;; ANTI-SENSE: NO
;; ORIGINAL SOURCE:
;; ORGANISM: Staphylococcus aureus
;; FEATURE:
;; NAME/KEY: misc-feature
;; LOCATION: 1...1119
US-08-827-356-1376

Query Match 32.8%; Score 416.2; DB 23; Length 1119;
Best Local Similarity 89.7%; Pred. No. 8e-83;
Matches 559; Conservative 0; Mismatches 44; Indels 20; Gaps 10;
OY 646 CTTATTAAACGATAAATTTCTCTTGATCTTAATTCATGATGATGATTTTGGT 705
DB 1 cttatttaacgataaattttctcc-ttgacatttaattcatgagttagga-ttttgg 58
OY 706 TTAATAGGTGATCATTTGAGTCAATTTATGATTCGATTCACCTCCAAATATAGT 765
DB 59 ttaataggtgatactttgagtcattttatgattatgcgattcaactgcaaatatagt 118
OY 766 ACAAGGTCAACAGGTGTTAATCCACCCGTTGGCGCTGTTGATGATGAAGTAGGAT 825
DB 119 acaaggtcaaacaggtgtaattccaccgcttgccgtgttagttantgaagtagagat 178
OY 826 TGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 885
DB 179 tgttggttattggttcacacttgagaaaggtgacaaagcatgagggttcaagcacttga 238
OY 886 TATGGCACAAANAAATGCTGAAGTGCAGCATTTATACGTAGAGCCATGTATGCA 945
DB 239 tatggcacaaanaaatgctgaagtgctgacgacatttatattacgttagaccatgtatgca 298
OY 946 TTTGTTCAACACACCCTGTGTTAACAAAATTATGATTGTAAGATAGCANAAAGTAGT 1005

Db	239	TTTTGGTTCACACACCACCCTGTGTACCAAAATTATTGATTGTAGATGCAAAAGTGT	358
OY	1006	ATTACNCAACANNAAGACAATTCGTTTAGACACACACATGGGTGATGAGACGTTACGGGGCTC	1065
Db	359	A-TACGCAACAAAAGACAATT-CGTTAGACACACAT-GGTGATGAGACGTTACGGGGCTCA	415
OY	1066	CACCGTATTTCGAGGTGCAATTGCGTTTGATATGTAAGCGGCATCACAATTATACCAG	1125
Db	416	CGGTATTGAGGTTCAATGCGCTT-----GATGATGAAACGGCGATCACAATTATPACCA--	466
OY	1126	ACTTTTATTAAAGCAAAAGCAACACTTGCACAAAATTATACAGTGAAGTNCCTTGA	1185
Db	467	AGACTTTTAAACCAAAAGCAAAAGCACT--CCACAATTACAGTGAAGTATTC-TGC	523
OY	1186	AAGTTAGATGGGTAACCAAGCGAATTGATTAATGACACAAAGTCAATGGATTACTAACAA	1245
Db	524	AAGTTAGATGGTAAACAAACG---ATGATATATGACAAAGTCATGGATTACTAACAA	580
OY	1246	AGAGGTTAAACAGATGCTATA	1268
Db	581	AGAGGTTAAACAGATGCTATA	603

```

RESULT      14
PCT-US98-24857-1
Sequence 1, Application PC/TUS9824857A
GENERAL INFORMATION:
APPLICANT: SmithKline Beecham Corporation
TITLE OF INVENTION: Novel rbg
FILE REFERENCE: P50444-9
CURRENT APPLICATION NUMBER: PCT/US98/24857A
CURRENT FILING DATE: 1998-11-23
EARLIER APPLICATION NUMBER: 08/978,456
NUMBER OF FILING DATE: 1997-11-25
NUMBER OF SEQ ID NOS: 7
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 1029
TYPE: DNA
ORGANISM: Staphylococcus aureus
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1026)
PCT-US98-24857-1

```

Query Match	27.8%;	Score 352.2;	DB 1;	Length 1029;
Best Local Similarity	88.4%;	Pred. No. 1.6e-68;		
Matches 473;	Conservative 0;	Mismatches 44;	Indels 18;	Gaps 8;

Oy	734	ATGATATATGGATTCAACTTCCAAATATGGTACAAAGGTTCANACAGGTGTATATCCACC	793
Db	1	atgattatgagattcaacttgcacaaatattgtcaaaagtctaaacagggttataccacc	60
Oy	794	GTTGGCGGTGTGTAGTTAATGAAGTAGAGATTGTGTATTGGTGACACTTGTGAAAA	853
Db	61	gttggcgctgtgttagttaaataagaagttagattgttgttattgggtgcacacttggaaaa	120
Oy	854	GGTGACAAGCATCGCGAAGTTCAAGCACTTGATATGGCAACACANANATCTGAAGTCCG	913
Db	121	gttgacaagcattgsggaggttcaagcacttgcattatgacacaaaatctgaaggtcg	180
Oy	914	ACGATTTATATTACGTTAGAGCAATGATCATTTGGTTCAAACACCACCCGTGTATAC	973
Db	181	acgatattatatacgttagagccacttgaagcatttgggttcaacacaccaccctgttatac	240
Oy	974	AAATATTATGATTCTTAAGATAGCANAGTAGTATGTTACNCAACANAAAGACATTCGGTTAG	1033
Db	241	aaattatttggtttgaagtatagcaaaagtgtat-tacgcacacaaagacaaatt-cgttag	298
Oy	1034	ACACACATGGTATGTAGACGTTACGGGGCTCCACGCGTATTGAGGCTTGAAATTCGTTG	1093

Db	299	acacacac	-gttga	tgaagc	gttaac	gggtcc	taagc	gtatga	gtttga	tgctt-----	351		
Qy	1094	GATGATGA	AGCGGCAT	CACTATT	ATACAA	AGACTTTT	TTTAA	GCACAAA	AAAGCAAA	GAA	1153		
Db	352	gatga	tgaagc	ggtatca	attata	cacga	actttt	---	aaag	caaa	gaagaa	408	
Qy	1154	CTTGCC	CAAAAT	TACAG	TAAAG	TCTTGA	AGTTTG	ATGG	TAAACA	AAAGCA	ATT	1213	
Db	409	ct--gc	caaca	attac	agtg	aaag	atc--tgc	agttga	tgtta	aaaca	agcgg--at	462	
Qy	1214	GATA	TGGACA	AAAGTCA	ATGGATT	TACTAT	CAAA	GAGGTAA	CAAGATG	CTAT	A	1268	
Db	463	gata	tgtca	aaagt	tcga	tgtatt	ctata	caaga	agttca	aaaca	agatgt	ctata	517

RESULT 15
US-09-376-633-1
: Sequence 1, Application US/09376633
: GENERAL INFORMATION:
: APPLICANT: Palmer, Leslie M.
: TITLE OF INVENTION: Novel ribg
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dechert Price & Rhoads
: STREET: 4000 Bell Atlantic Tower, 1717 Arch Street
: CITY: Philadelphia
: STATE: PA
: COUNTRY: US
: ZIP: 19103
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Discrete

```

1  COMPUTER: IBM Compatible
2  OPERATING SYSTEM: DOS
3  SOFTWARE: FASTSEQ for Windows Version 1.0
4  CURRENT APPLICATION DATA:
5  APPLICATION NUMBER: US/09/376,633
6  FILING DATE:
7  CLASSIFICATION:
8  PRIOR APPLICATION DATA:
9  APPLICATION NUMBER: 08/978,456
10 FILING DATE:
11 ATTORNEY/AGENT INFORMATION:
12 NAME: Dickinson, Todd Q
13 REGISTRATION NUMBER: 28,354
14 REFERENCE/DOCKET NUMBER: P50444-9
15 TELECOMMUNICATION INFORMATION:
16 TELEPHONE: 215-994-2252
17 TELEFAX: 215-994-2222
18 TELETYPE:
19 INFORMATION FOR SEQ ID NO: 1:
20 SEQUENCE CHARACTERISTICS:
21 LENGTH: 1029 base pairs
22 TYPE: nucleic acid
23 STRANDEDNESS: double
24 TOPOLOGY: linear
25
26 US-09-376-633-1

```

Query Match	27.8%;	Score 352.2;	DB 103;	Length 1029;
Best Local Similarity	88.4%;	Pred. No. 1.6e-68;		
Matches 473; Conservative	0;	Mismatches 44;	Indels 18;	Gaps 8

OY	734	ATGATATATGCGATTCAACTTCCAAATATGGTACAAAGTCAMCAAGGTGTTAATCCACC	733
Db	1	ATGATATATGCGATTCAACTTCCAAATATGGTACAAAGTCAMCAAGGTGTTAATCCACC	60
OY	794	GTTCGCGTGTGTAGTTAATGAAGGTAGCATTTTGTGTAATGGTGACACTTGAGAAA	853
Db	61	GTTCGCGTGTGTAGTTAATGAAGGTAGCATTTTGTGTAATGGTGACACTTGAGAAA	120
OY	854	GGTGCAAGCATTCGGGAGGTTCACAGCACTTGATATGGCACACAMCAATGCTGAAGGTGC	913.3
Db	121	GGTGCAAGCATTCGGGAGGTTCACAGCACTTGATATGGCACACAMCAATGCTGAAGGTGC	180

```

QY 914 ACGATTATATACGTTAGACCATGATGATTTGGTCCAAACACCCTGTGTTAC 973
    |||||||
DB 181 ACGATTATATATACGTTAGACCATGATGATTTGGTCCAAACACCCTGTGTTAC 240
    |||||||
QY 974 AAAATTATGATGTAGATAGCANAAGTAGTATTACNCAACANAAGACAATTCCGTTAG 1033
    |||||||
DB 241 AAAATTATGATGTAGATAGCANAAGTAGTATTACNCAACANAAGACAATT-CGTTAG 298
    |||||||
QY 1034 ACACACATGGGTGATGAGACGTTACGGGGCTCCACGGTATTGAGGGTGAATTGCGTTG 1093
    |||||||
DB 299 ACACACAT-GGTGATGAGACGTTACGGGGCTCCACGGTATTGAGGGTGAATGCGTT----- 351
    |||||||
QY 1094 GATGATGAAGGGCATCACATTTATACCAAGACTTTTAAAGCAAAAAGCAAGCAA 1153
    |||||||
DB 352 GATGATGAAGGGCATCACATTTATACCAAGACTTTT--AAAGCAAAAAGCAAGCAA 408
    |||||||
QY 1154 CTGGCCAAATTTACAGTGAAGTCTGAAAGTTTAGATGGGTAACCAAGCGAAT 1213
    |||
DB 409 CT--GCCACAAATTTACAGTGAAGTATC-TGCAAGTTTAGATGGTAAACAAAGCG--AAT 462
    |||
QY 1214 GATTAATGGACAAAGTCAATGATTTACTTAACAAAGAGGTTAAACAGATGCTATA 1268
    |||||||
DB 463 GATTAATGGACAAAGTCAATGATTTACTTAACAAAGAGGTTAAACAGATGCTATA 517
    |||||||

```

Search completed: November 19, 2000, 05:57:52
 Job time: 8499 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 19, 2000, 00:50:17 ; Search time 2632.38 seconds
(without alignments)
1707.291 Million cell updates/sec

Title: US-08-978-456-1
Perfect score: 1029
Sequence: 1 ATGGATTATGCGATTCAACT.....TAAGTTACGAAGAAGTGA 1029

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1033670 seqs, 2183789903 residues

Total number of hits satisfying chosen parameters: 2067340

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pl1:*
10: gb_pr2:*
11: gb_pr3:*
12: em_fun:*
13: em_hum1:*
14: em_hum2:*
15: em_in:*
16: em_om:*
17: em_or:*
18: em_ov:*
19: em_pat:*
20: em_ph:*
21: em_pl:*
22: em_ro:*
23: em_sts:*
24: em_sy:*
25: em_un:*
26: em_v1:*
27: gb_hcg1:*
28: gb_hcg2:*
29: gb_in1:*
30: gb_in2:*
31: em_ba1:*
32: em_ba2:*
33: em_hum3:*
34: em_hum4:*
35: gb_pr4:*
36: gb_hcg3:*
37: gb_hcg4:*
38: gb_hcg5:*
39: gb_hcg6:*
40: gb_hcg7:*
41: em_hcg1:*
42: em_hcg2:*
43: em_hcg3:*

44: em_hum5:*
45: gb_pl3:*
46: gb_pr5:*
47: gb_hcg8:*
48: gb_hcg9:*
49: gb_hcg10:*
50: gb_hcg11:*
51: gb_hcg12:*
52: gb_hcg13:*
53: gb_hcg14:*
54: gb_in3:*
55: gb_hcg15:*
56: gb_hcg16:*
57: gb_hcg17:*
58: em_hcg4:*
59: em_hcg5:*
60: em_hcg6:*
61: em_hcg7:*
62: em_hum6:*
63: gb_hcg18:*
64: gb_hcg19:*
65: gb_ba3:*
66: em_hcg8:*
67: em_hcg9:*
68: em_hcg10:*
69: em_hcg11:*
70: em_hcg12:*
71: em_hcg13:*
72: em_hcg14:*
73: em_hcg15:*
74: em_hcg16:*
75: em_hcg17:*
76: em_hcg18:*
77: em_hcg19:*
78: em_hcg20:*
79: em_hcg21:*
80: em_hcg22:*
81: em_hcg23:*
82: gb_pr6:*
83: gb_pr7:*
84: gb_hcg20:*
85: gb_hcg21:*
86: gb_hcg22:*
87: gb_hcg23:*
88: gb_ro:*
89: gb_sts1:*
90: gb_sts2:*
91: gb_sy:*
92: gb_un:*
93: gb_v11:*
94: gb_v12:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1029	100.0	3335	AR069579	AR069579-sequence--
2	516.8	50.2	3227	AF269646	AF269646 Staphyloc
3	516.8	50.2	3618	AF269345	AF269345 Staphyloc
4	340.6	33.1	2802	AF270130	AF270130 Staphyloc
5	155	15.1	3721	AF269926	AF269926 Staphyloc
6	120.4	11.7	6006	BSRIB	X51510 B.subtilis
7	120.4	11.7	28206	BACDIA	L09228 Bacillus su
8	120.4	11.3	218470	BSUB0013	299116 Bacillus su
9	116.4	11.3	14364	AE000675	AE000675 Aquifex a
10	112.2	10.9	4286	BARIBGENS	X95955 B.amycoliqu
11	95.2	9.3	18073	U32775	U32775 Haemophilus
12	91.4	8.9	11695	AE004298	AE004298 Vibrio ch

Db	61	GTTCGGCCCTTCTTAACTTATGACAGTGAAGATGTTGGTATTGGTGCACACTTGGAGAAA	120
QY	121	GGTGACAAAGCATGCGAGGTTCCAAAGCCTGATGATGCGACAAACAAAATGCTGAAGGTGG	180
Db	121	GGTGACAAAGCATGCGAGGTTCCAAAGCCTGATGATGCGACAAACAAAATGCTGAAGGTGG	180
QY	181	ACGATTATATTCGTTAGGACGATGTGTCATTTGGTTTAAACACACCTGGTTAAC	240
Db	181	ACGATTATATTCGTTAGGACGATGTGTCATTTGGTTTAAACACACCTGGTTAAC	240
QY	241	AAATATTGATGTGTAGATAGCAAAAGTAGATATACGCACACAAAAGCAATTGCTTAGAC	300
Db	241	AAATATTGATGTGTAGATAGCAAAAGTAGATATACGCACACAAAAGCAATTGCTTAGAC	300
QY	301	ACACATGGTATATAGACGTTAGGGGCTACGGTATGAGGTTGAATCGCTGATGATGAA	360
Db	301	ACACATGGTATATAGACGTTAGGGGCTACGGTATGAGGTTGAATCGCTGATGATGAA	360
QY	361	CGGGCATCACAAATTATATCCAAAGACATTTTTTAAAGCAAAAGCAAAAGCAACTGCCAAT	420
Db	361	CGGGCATCACAAATTATATCCAAAGACATTTTTTAAAGCAAAAGCAAAAGCAACTGCCAAT	420
QY	421	ACAGTGAATATCTGCACAGTTTAAATGCTAAACACGGAATGATATGGACAAAGTCAA	480
Db	421	ACAGTGAATATCTGCACAGTTTAAATGCTAAACACGGAATGATATGGACAAAGTCAA	480
QY	481	TGCAATTACTAACAAAGGTTTAAACAAAGTCTTATAAGTTAAAGACATGCACAGACGA	540
Db	481	TGCAATTACTAACAAAGGTTTAAACAAAGTCTTATAAGTTAAAGACATGCACAGACGA	540
QY	541	GTGTAACTGGAAGACGTACAGTTGAATTAGATGATCCACAAATACTACACGATTTCAA	600
Db	541	GTGTAACTGGAAGACGTACAGTTGAATTAGATGATCCACAAATACTACACGATTTCAA	600

OY	601	GATGAAAAAACCCCTTAATAAGTATATGTCCTAAGTGGGAATATTCATTTTAAACAG	660
Db	601	GATGAAAAAACCCCTTAATAAGTATATGTCCTAAGTGGGAATATTCATTTTAAACAG	660
OY	661	CAAAATTATCAAGATGAATCAACACCAAATTTGGATATATACTGAAAAATCCAAATTTTACA	720
Db	661	CAAAATTATCAAGATGAATCAACACCAAATTTGGATATATACTGAAAAATCCAAATTTTACA	720
OY	721	AGCAATCAAAACACATATTTGAAATTTATTTACTTGAAGCTTGATTTAACAACAAATTCCT	780
Db	721	AGCAATCAAAACACATATTTGAAATTTATTTACTTGAAGCTTGATTTAACAACAAATTCCT	780
OY	781	CACATTTTATATAAAGAGATGTGGAACTTTCGACTGCGAGGACAGGTCCACCACTACT	840
Db	781	CACATTTTATATAAAGAGAGATGTGGAACTTTCGACTGCGAGGACAGGTCCACCACTACT	840
OY	841	TCGAATTCCTCATATATATAGATGAATTTATTCCTATTTAGCCCGCAAAATTAAAT	900
Db	841	TCGAATTCCTCATATATATAGATGAATTTATTCCTATTTAGCCCGCAAAATTAAAT	900
OY	901	GGCGGATCTGGAATTTATCAATTTTATCAAAACAAATGATGTGATTGAGATACAGATGCG	960
Db	901	GGCGGATCTGGAATTTATCAATTTTATCAAAACAAATGATGTGATTGAGATACAGATGCG	960
OY	961	AACCAATTTGAAATTTGTTATTTCCAGATTATTAATCAAATGTTAAATTAACCTTTACGA	1020
Db	961	AACCAATTTGAAATTTGTTATTTCCAGATTATTAATCAAATGTTAAATTAACCTTTACGA	1020

Db	1021	AAAGAGTGA	1029
RESULT	2		
AF269646/c	AF269646	3227 bp	DNA
LOCUS	Staphylococcus epidermidis strain Sfi1 clone	step.1015d10	genomic
DEFINITION	sequence.		

ACCESSION AF269646 GI:9623542
 VERSION AF269646.1
 KEYWORDS Staphylococcus epidermidis.
 SOURCE Staphylococcus epidermidis.
 ORGANISM Bacillus/Clostridium group; Staphylococcus.
 REFERENCE 1 (bases 1 to 3227)
 AUTHORS Kimmerly, W.J., Taylor, J. David, Nelson, A.J., Godlevski, M.M.,
 Rubino, M.A., Nelson, F.J., Rivers, P.R., Torruella-Miller, I.,
 Listenebee, S., Ashanti, C., Altschuller, G., Mamou, L., Shepherd, N.S.,
 Fuchs, R., Fleming, T., Guan, X., Du, L., Cain, D.H., Miller, G.S. and
 Furdon, P.J.
 TITLE Transposon-mediated sequencing of the Staphylococcus epidermidis
 genome
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 3227)
 AUTHORS Taylor, J. David, Kimmerly, W.J., Nelson, A.J., Godlevski, M.M.,
 Rubino, M.A., Nelson, F.J., Rivers, P.R., Torruella-Miller, I.,
 Listenebee, S., Ashanti, C., Altschuller, G., Mamou, L., Shepherd, N.S.,
 Fuchs, R., Fleming, T., Guan, X., Du, L., Cain, D.H., Miller, G.S. and
 Furdon, P.J.
 TITLE Direct Submission
 JOURNAL Submitted (22-May-2000) Departments of Genomic Sciences and
 Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore
 Drive, Research Triangle Park, North Carolina 27709-3398, USA
 FEATURES
 source
 1..3227
 /organism="Staphylococcus epidermidis"
 /strain="SRI"
 /db_xref="taxon:1282"
 /clone="step.1015d10"
 BASE COUNT 980 a 599 c 454 g 1194 t
 ORIGIN
 Query Match 50.2% Score 516.8; DB 1; Length 3227;
 Best Local Similarity 69.5% Pred. No. 3.4e-85;
 Matches 717; Conservative 0; Mismatches 312; Indels 3; Gaps 1;

Db 1438 TGGATTACCAACAAAGAACTTAAGAGATGTTTCAATTAGACATAGCATGATGCA 1379
 QY 541 GTGTTACTGGAAGAGCTACAGTTGATTAAGATCCCAATATACATACAGTATTCGA 600
 Db 1378 GTTATTACTGGCGTGAAGAACCAATTGAAGCAACATCTTGTATACCAACGAGGCTT 1319
 QY 601 GATGAAAAAACCTTATTAAGTAATATGTTAGTGGGAAATATTCATTTTAATGAG 660
 Db 1318 GATGAAACACATCCGATTCGATGATTTCTTTAGAAAGTCACACTGATTTTAATCA 1259
 QY 661 CAAATTTATCAAGATGAATCAACACCAATTTGATATATCTGAAATCAAAATTTACA 720
 Db 1258 CAAATTTTAAGATACCTGATCGAGATGATGATTTACATGAAAAATTAATAA 1199
 QY 721 AGCATCAACACATTTTAATATTTATTTCTGGAAGTCTTGATTTTACCAATATCT 780
 Db 1198 ACAAAATAAAGTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1139
 QY 781 CCAATTTTATTAAGAGAGAGTGGAGCTTGTAGTGGAGAGTGGAGTGGAGTGGAGTGG 840
 Db 1138 CAAGACTTATATCAAGAGAGTATGGGAACCTGTAGTGGAGGCAAGCCCAATATTA 1079
 QY 841 TCAGATTTCTCCATCT--ATTATATAGATGAATTTATCTATATTATGCCCCGAATTA 897
 Db 1078 TCTCAATTTCTCCATCTCAACATCAATCAATCAATCAATCAATCAATCAATCAATTA 1019
 QY 898 ATTGGCGGATCTGGAATTTATCAATTTTCAACAAATGATGATTAATACAGAT 957
 Db 1018 ATTGGGTTCTGCGCAACATCAATTTTAAAGACGAGAGTGAATTAATTTGCTGA 959
 QY 958 GCGAACCAATTTGAATTTGTCATTCGATTCGATTTAATCAAAATGTTAATTAATTA 1017
 Db 958 GCACTCAATTTGAATTTGTCATTCGATTTAATCAAAATGTTAATTAATTAATTA 899
 QY 1018 CGAAGAGCTGA 1029
 Db 898 CGAAGAGCTGA 887
 RESULT 3
 AF269345/c 3618 bp DNA BCT 03-AUG-2000
 LOCUS Staphylococcus epidermidis strain SRI clone step.1002d07 genomic
 DEFINITION
 accession AF269345
 version AF269345.1 GI:9664666
 keywords
 source Staphylococcus epidermidis.
 organism Staphylococcus epidermidis.
 reference Bacillus/Firmicutes: Bacillus/Clostridium group;
 authors Kimmerly, W.J., Taylor, J. David, Nelson, A.J., Godlevski, M.M.,
 Rubino, M.A., Nelson, F.J., Rivers, P.R., Torruella-Miller, I.,
 Listenebee, S., Ashanti, C., Altschuller, G., Mamou, L., Shepherd, N.S.,
 Fuchs, R., Fleming, T., Guan, X., Du, L., Cain, D.H., Miller, G.S. and
 Furdon, P.J.
 title Transposon-mediated sequencing of the Staphylococcus epidermidis
 genome
 journal Unpublished
 reference 2 (bases 1 to 3618)
 authors Taylor, J. David, Kimmerly, W.J., Nelson, A.J., Godlevski, M.M.,
 Rubino, M.A., Nelson, F.J., Rivers, P.R., Torruella-Miller, I.,
 Listenebee, S., Ashanti, C., Altschuller, G., Mamou, L., Shepherd, N.S.,
 Fuchs, R., Fleming, T., Guan, X., Du, L., Cain, D.H., Miller, G.S. and
 Furdon, P.J.
 title Direct Submission
 journal Submitted (19-May-2000) Departments of Genomic Sciences and
 Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore
 Drive, Research Triangle Park, North Carolina 27709-3398, USA
 features
 source 1..3618

```
/organism="Staphylococcus epidermidis"
/strain="SRI"
/db_xref="taxon:1282"
/clone="step.1002d07"
BASE COUNT      1099 a      709 c      518 g      1292 t
ORIGIN

Query Match      50.2%  Score 516.8;  DB 1;  Length 3618;
Best Local Similarity 69.5%;  Pred. No. 3.3e-85;
Matches 717;  Conservative 0;  Mismatches 312;  Indels 3;  Gaps 1;

OY 1 ATGGATTATGCGATTCACCTGCAATATGTTACAGCTCAACAGCTGTTAATCCACC 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2701 ATGGATGATGCTATTCACATGCAAGAAATGGAATGAGCAACAGCTTATCCACA 2642
OY 61 GTTGGCGCTGTTAGTAATGAAGTAGAATGTTGTTGTTGTTGTTGTTGTTGTTA 120
    || || || || || || || || || || || || || || || || || || || || ||
Db 2641 GTAGATCCGTTGTTGTTAAAAACGCTAGCATTTAGCTTATGCTGACATTTAAAAG 2582
OY 121 GGTGACAAAGCATCGGAGCTCAAGCACTTATATGACACAAACAAATGCTGAAGTGG 180
    || || || || || || || || || || || || || || || || || || || || ||
Db 2381 GGACATTAACATGCGGAAGTCAACGCTATGAAATGGAGTTTAAATACCCAAAGTCT 2522
OY 181 AGAATTTATATGAGTTCAGCATGTATGCTATTTGTTCAACACACCCCTGTTTAA 240
    || || || || || || || || || || || || || || || || || || || || ||
Db 2521 ACCATATACGTTTCTTATGAACTTGCACACACCCCTGTTCAACCCCTTGTGCTAT 2462
OY 241 AAAATTTATGATTGTAGATGACAAAGATAGTATACGACAAAGAACATTCGTTAGC 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2461 AAAATCATTTGAAGCGGCGCATCTCTAGTCTCTATGCTGTTAAAGATCTACTTACTA 2402
OY 301 ACACATGCTGATGAGACCTTACGGGCTACGGTATGAGTTGATGCTGATGATGAA 360
    || || || || || || || || || || || || || || || || || || || || ||
Db 2401 AGTAGGGTGCAGATTCTAGAGAACCTGCTATAGAGGTTGAATTTCAATATATATAA 2342
OY 361 CGGCATCATATATATACCAAGACTTTTAAAGCAAGCAAGCAAGCTGCCAATTT 420
    || || || || || || || || || || || || || || || || || || || || ||
Db 2341 AATGACGCTGATTTATACCGTGACCTTTTCTACTGCTAAAGCAAGCAAGCTGCCAAGTA 2282
OY 421 ACAGTGAAGATATGCAAGTTAGATGTTAAGATGAACAGCAAGCAAGTAAATGACAAAGTCAA 480
    || || || || || || || || || || || || || || || || || || || || ||
Db 2281 ACTGTAAGAGTCTCATCTAGTCTAGATGTTAACAAGCAAGCAAGCTTAAATGAAGTAAG 2222
OY 481 TGGATTACTAACAAGAGGTTAAACAAATGCTTAAAGATGAACATCGACAGCAAGCA 540
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2221 TGGATTACAAACAAGAGTTAAAGAAAGATGTTTATCAATTAAGCATGAGCATGATGATCA 2162
OY 541 GTGTTACTGGAAGACGATGAGTGAATGATGATCCACATATACTACAGTATTCAA 600
    || || || || || || || || || || || || || || || || || || || || ||
Db 2161 GTTATTACTGGCGCTAGAACCATTTGAAGCAGACATCCATTTATACACAGGTTCTT 2102
OY 601 GATGAAAAAACCTATTAAGATATGTTCTAAGTCTGGCAATATTCATTTTAATCAG 660
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2101 GATGAAAGCATCGGATTCGAGTATATCTTCTTAAGAAAGGTCAACTGATTTTAATCAA 2042
OY 661 CAATTTATCAAGATGATCAACACCAATTTGATATATCTGAAATTCCAATTTAACA 720
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2041 CAATATATTTAAATGATCTGACATCGAGATGATTTACCTGAAATGAAGAAATTTAAA 1982
OY 721 AGCATCAACACATATGATTAATTTACTTGAAGCTTGTGATTTAACAACAATTTCT 780
    || || || || || || || || || || || || || || || || || || || || ||
Db 1981 ACAATATTAATTTATTAATTAATTAATTAATTAATTTGATTTGATACAGACATTAATTA 1922
OY 781 CACAATTTATTAAGAGAGGATTTGGAATTTGCTAGTTCGAGGCGAGTTCACCACTACT 840
    || || || || || || || || || || || || || || || || || || || || ||
Db 1921 CAAGACTATATCAAGAGGATTTGGGAACTGCTAGTTCGAGGCGAGGCGCAATATTAACA 1862
OY 841 TCAGATTTCTCACTCT--ATTATATAGATGAATTTATCTCTATTTATGCCCGGAATTA 897
    || || || || || || || || || || || || || || || || || || || || ||
Db 1861 TCTCATATTTCTCCATCCAAACATCTTAATGACATTTTATATATATGATGATGATGATTA 1802
OY 898 ATTGGCGATCTGAAATTTATCAATTTTATCAACAATATGATGATGATGATGATGAT 957
```

```
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1801 ATTGGGTGCTCTGGCAACATCATTTTATTAAGACTGACAGGCTGATTTGCTGAA 1742
OY 958 GCGAACCAATTTGAAATTTGTTTCATTCGAGTATTTAATTAATGAAATGTAATTAATTA 1017
    || || || || || || || || || || || || || || || || || || || || ||
Db 1741 GCAACTCAATTTGAAATTTGTTGATTCGAGTATTAATTAATTAATTAATTAATTA 1682
OY 1018 CGAAGAAGATGA 1029
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1681 CGAAGAAGATGA 1670

RESULT 4
AF270130 2902 bp DNA BCT 01-AUG-2000
LOCUS Staphylococcus epidermidis strain SRI clone step.1051a10 genomic
DEFINITION sequence.
ACCESSION AF270130
VERSION AF270130.1 GI:9624037
SOURCE Staphylococcus epidermidis.
ORGANISM Staphylococcus epidermidis.
Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Staphylococcus.
1 (bases 1 to 2902)
REFERENCE Kimmerly,W.J., Taylor,J.David, Nelsen,A.J., Godlevski,M.M.,
AUTHORS Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
Listenbee,S., Ashanti,C., Altschuller,G., Mamou,L., Shepherd,N.S.,
Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
Furdon,P.J.
Transposon-mediated sequencing of the staphylococcus epidermidis
genome
2 (bases 1 to 2902)
REFERENCE Taylor,J.David, Kimmerly,W.J., Nelsen,A.J., Godlevski,M.M.,
AUTHORS Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
Listenbee,S., Ashanti,C., Altschuller,G., Mamou,L., Shepherd,N.S.,
Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
Furdon,P.J.
Direct Submission
Submitted (22-MAY-2000) Departments of Genomic Sciences and
Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore
Drive, Research Triangle Park, North Carolina 27709-3398, USA
FEATURES
source
1..2902
/organism="Staphylococcus epidermidis"
/strain="SRI"
/db_xref="taxon:1282"
/clone="step.1051a10"
BASE COUNT 1009 a 382 c 507 g 1004 t
ORIGIN

Query Match 33.1%; Score 340.6; DB 1; Length 2902;
Best Local Similarity 69.0%; Pred. No. 5.1e-53;
Matches 466; Conservative 0; Mismatches 209; Indels 0; Gaps 0;

OY 1 ATGGATTATGCGATTCACCTGCAATATGTTACAGCTCAACAGCTGTTAATCCACC 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2220 ATGGATGATGCTATTCACATGCAAGAAATGGAATGAGCAACAGCTTATCCACA 2279
OY 61 GTTGGCGCTGTTAGTAATGAAGTAGAATGTTGTTGTTGTTGTTGTTGTTGTTA 120
    || || || || || || || || || || || || || || || || || || || || ||
Db 2280 GTAGATCCGTTGTTGTTAAAAACGCTAGCATTTAGCTTATGCTGACATTTAAAAG 2339
OY 121 GGTGACAAAGCATCGGAGCTCAAGCACTTATGATGCAACAAATGCTGAAGTGG 180
    || || || || || || || || || || || || || || || || || || || || ||
Db 2340 GGAGATTAACATGCGGAAGTCAACGCTATTTGAATGATGCTTAAATACCAAGTGTCT 2399
OY 181 AGAATTTATATGAGTTCAGCATGTATGCTATTTGTTGTTCAACACCCCTGTGTTAC 240
    || || || || || || || || || || || || || || || || || || || || ||
Db 2400 ACCATATACGTTTCATTAAGAACCTTGCACACACCATGTTCAACACACCTGTGTGCTAT 2459
```

OY	241	AAAAATATTGATGTGAAGTATACCAAAAGTAGTAATACGGACAACAAAAGACAAATTCGTATAC	300
Dd	2460	AAAATCATTTGAAGCGGGGCATATCTAAGGTCATCTATGCTTTAAAGABACTACTTTAGTA	2519
OY	301	ACACATGGTGATGATGACGCTTTACGGGCTCCAGGTATTTGAAGTTGAATGCCGTGATGATA	360
Dd	2520	AGTAAAGGATGACGAGAATTTCTGAGAGAGCGTGGTATAGAGGCTGAATTTCAATATATATGA	2579
OY	361	CGGGCATCACAATTTATACCAGACTTTTTTTTAAAGCAAAGCAAGCAACTGCCACAAATT	420
Dd	2580	AATGACAGCTGCATTTTATACCGTGACTCTTTTACTGCTAAAAGAAAGCAAGTTCCACAAGTA	2639
OY	421	ACAGTAAAGATATTCGCAAGTTTACGATTTAGTGGTAAACAGCAAGCAATATATGACAAAGTCA	480
Dd	2640	ACTGTAAAGGTCATCTATGATCTAGATGTAAACACAGCACAGACTTTTATATAAAGTAAAG	2699
OY	481	TGGATTACTAACAAAGAGGTTTAAACAGATGCTATAAGTTAAAGACATCGACAGACGCA	540
Dd	2700	TGGATTAACAACAAAGAAAGTTAAAGAGATGTTTATCATTTAAGACATGAGCATATGCA	2759
OY	541	GCGTTAACTGGAAGACGTACACTTGAATTAAGTATGATGATCCCAATATATACAGTATTCGA	600
Dd	2760	GTTATTACTGGGCGGTAGAACCATTGAAGCAGCAATTCATTTGATATCAACCGGGTTCT	2819
OY	601	GATGCAAAAAACCCCTATATAAAGTAATATATGTCTTAAGTCTGGGAATATTCATTTAATCAG	660
Dd	2820	CATGAAAGACATCCGATTCGAGTTATCTTTCTTAAAGAAAGTGCAACTGATTTTATATCA	2879
OY	661	CAAAATTTATCAGAT	675
Dd	2880	CAAAATTTTAAAGAT	2894
RESULT	5		
AF269926/c			
LOCUS	AF269926	3721 bp	DNA
DEFINITION	Staphylococcus epidermidis strain sri clone step.1032d02 genomic sequence.		
ACCESSION	AF269926		
VERSION	AF269926.1	GI:9623827	
KEYWORDS			
SOURCE	Staphylococcus epidermidis.		
ORGANISM	staphylococcus epidermidis		
REFERENCE	Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group; Staphylococcus.		
AUTHORS	1 (bases 1 to 3721)		
	Kimberly,W.J., Taylor,J.David, Nelsen,A.J., Godlewski,M.M.,		
	Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,		
	Listebeebe,S., Ashanti,C., Althuller,G., Memo,L., Shepherd,N.S.,		
	Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and		
	Furdon,P.J.		
TITLE	Transposon-mediated sequencing of the Staphylococcus epidermidis genome		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 3721)		
AUTHORS	Taylor,J.David, Kimberly,W.J., Nelsen,A.J., Godlewski,M.M.,		
	Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,		
	Listenbee,S., Ashanti,C., Althuller,G., Memo,L., Shepherd,N.S.,		
	Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and		
	Furdon,P.J.		
TITLE	Direct Submission		
JOURNAL	Submitted (22-MAY-2000) Departments of Genomic Sciences and		
	Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore		
	Drive, Research Triangle Park, North Carolina 27709-3398, USA		
FEATURES	Location/Qualifiers		
source	1..3721		
	/organism="Staphylococcus epidermidis"		
	/strain="SRI"		
	/db_xref="taxon:1282"		
	/clone="step.1032d02"		
BASE COUNT	1204 a	625 c	495 g 1397 t
ORIGIN			

Query Match	15.1%	Score 155	DB 1	Length 3721
Best Local Similarity	76.1%	Pred. No. 3.8e-19		
Matches 191	Conservative 0	Mismatches 60	Indels 0	Gaps 0
QY	1	ATGATATTCGATTCACACTTGCAAAATATGGTACAGGTCAACAGAGTGTAAATCCACC	60	
Db	259	ATGATGATTCGATTCACACTAGCAAAATATGGTAAATGACAAACAGGTGTAAATCCACCA	200	
QY	61	GTTGGCGCTGTTTGTAAGTAAATGAAGTATGGATGTGTGTTATTTGGTCACACTTGAGAAA	120	
Db	199	GTAGATATCCGTTGTTGTTAAAAACGGTAGATGTGATGGTTTAGGTCACACTTTAAAAAG	140	
QY	121	GGTGACAGCATGCGGAGGTCACAGCACTGTATATAGCACAAACAAATGCTGAAGTGC	180	
Db	139	GGAGTAAACATGCGCGAGTAGCAAGCTATTGAAATGCGAGGTTTAAATACCCAGAGTCT	80	
QY	181	ACGATTATATTTACGCTTAGACCATGATGATCTATTTGGTTCAACACACCACTGTGTAC	240	
Db	79	ACCAATATTCGTTTCATTATGAACTTGACACACACATGCTTCAACACCACTTGTTGCAT	20	
QY	241	AAATTTATGA	251	
Db	19	AAATCATTGA	9	
RESULT	6			
LOCUS	BSR1B	6006 bp	DNA	BCT
DEFINITION	B.subtilis riboflavin biosynthesis operon ribG, ribB, ribA, ribH, and ribT genes.			10-FEB-1999
ACCESSION	X51510			
VERSION	X51510.1	GI:40083		
KEYWORDS	ribA gene; ribB gene; ribG gene; ribH gene; riboflavin biosynthesis; riboflavin biosynthesis operon; ribT gene.			
SOURCE	Bacillus subtilis.			
ORGANISM	Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillaceae; Bacillus.			
REFERENCE	1 (bases 1 to 6006)			
AUTHORS	Mironov,V.N.			
JOURNAL	Direct Submission			
REMARK	Submitted (23-JAN-1990) Mironov V.N., Centre of Bioengineering, Institute of Molecular Biology, Academy of Science, Vavilov str 32, Moscow 117984, U S S R (AA 1-6006)			
REFERENCE	2 (bases 858 to 1246)			
AUTHORS	Mironov,V.N., Perumov,D.A., Kraev,A.S., Stepanov,A.I. and Sriadin,K.G.			
TITLE	Unusual structure of the regulatory region of the riboflavin biosynthesis operon in Bacillus subtilis			
JOURNAL	Mol. Biol. (Mosk.)	24 (1),	256-261 (1990)	
REMARK	90271920			
REMARK	(AA 1-6006)			
REMARK	3 (bases 1 to 6006)			
REMARK	K.I.I.Y.V., Mironov,V.N., Gorishin,I.Yu., Krenева,R.A. and Perumov,D.A.			
TITLE	Riboflavin operon of Bacillus subtilis: unusual symmetric arrangement of the regulatory region			
JOURNAL	Mol. Gen. Genet.	233 (3),	483-486 (1992)	
COMMENT	92318904			
FEATURES	Data kindly reviewed (03-SEP-1990) by Mironov V.N.			
SOURCE	Location/Qualifiers			
	1..6006			
	/organism="Bacillus subtilis"			
	/strain="SHGW"			
	/db_xref="taxon:1423"			
	/map="210 units"			
	475..819			
	/note="ORF X (AA 1-114)"			
	/codon_start=1			
	/transl_table=11			
	/protein_id="CAA35877.1"			

SOURCE Bacillus subtilis (strain 168, sub-species Marburg) DNA.
ORGANISM Bacillus subtilis
Bacteria: Firmicutes: Bacillus/Clostridium group:
Bacillus/Staphylococcus group; Bacillus.
REFERENCE 1 (bases 1 to 28206)
AUTHORS Yamamoto,J., Shimizu,M. and Yamane,K.
TITLE Molecular cloning and analysis of nucleotide sequence of the
Bacillus subtilis lysA gene region using B. subtilis phage vectors
and a multi-copy plasmid. PubMed
JOURNAL Agric. Biol. Chem. 55 (6), 1615-1626 (1991)
MEDLINE 91345841
REFERENCE 2 (bases 1 to 28206)
AUTHORS Buchanan,C.E. and Ling,M.L.
TITLE Isolation and sequence analysis of dacB, which encodes a
sporulation-specific penicillin-binding protein in Bacillus
subtilis
JOURNAL J. Bacteriol. 174 (6), 1717-1725 (1992)
MEDLINE 92193254
REFERENCE 3 (bases 1 to 28206)
AUTHORS Sorokin,A., Zumbstein,E., Azavedo,V., Ehrlich,S.D. and Serrro,P.
TITLE The organization of the Bacillus subtilis 168 chromosome region
between the spoVA and serA genetic loci, based on sequence data
JOURNAL Mol. Microbiol. 10 (2), 385-395 (1993)
MEDLINE 95020538
FEATURES
source location/Qualifiers
1. 28206
/organism="Bacillus subtilis"
/strain="168"
/sub_species="Marburg"
/db_xref="taxon:1423"
1. 1239
/gene="spoVAF"
<1. 1239
/gene="spoVAF"
/codon_start=1
/transl_table=11
/protein_id="AA67472.1"
/db_xref="GI:410115"
/translation="VDIVENRLMAOVKVKLTDETTDOVLGSLAVIVYEGAFATII
DVSRYGRNPEEDPTKEVYRGARGGVENVYVNTALLRRIRDERLRVAKTKYGEKSK
TDLSTICYIEDIADPDLVEYKFIASIDVGLTMAKIVAEETLVNOSYNPFLVRYTE
RDPVANHVLGHHVHIIIVDTSPEVITPTLTFHHVHAEYRQAPSGTFLRWVRFG
IASTLFLPWFLEPDLDPNMKFIQKNTQTHPILOIFLADLGEFLMAAHS
TPTALSTAMGLAAVLIGQIAEVLGFEFSEVILVYSLAIGFTTPTSYELSTATNPS
CHRDTRCFISYKARHRRLKSANMAASINSLOPIYMLPIFNGKALNOVLVYRTAKPS
AAVRSTIVHPKRLRQPTNS"
1158. 1164
/gene="spoVAF"
/citation=[1]
1245. 1250
/gene="lysA"
/note="putative"
1245. 2664
/gene="lysA"
1266. 1272
/gene="lysA"
/note="putative"
1345. 2664
/gene="lysA"
/codon_start=1
/transl_table=11
/product="diaminopimelate decarboxylase"
/protein_id="AA67473.1"
/db_xref="GI:410117"
/translation="MFLHGTSRQNHGHLEIGVDALYLAERYGPLYVYDVALLER
AKSFOAFISAGLKAQVAAASKAFSSVAMQLAEEGSLDVAVSGELLYTAAAGPPA
ERIHFGNNKSRLELMALHRIQCIYVDNFIALLDLCKETGSHSIDVLLRITPGV
EATHTHYITTGDSKFGFDLHNGOTERAEIOVLOSEHIOLGHHCHISQJEDITGF
VLAERKIFKKLDWMRSYSFYSKVLNLGGFCGRYEDDEPRLHATYVKKITEAVKEN
ASRYGFDIPEIWIIEGRSLVGDAQTLLIYTGSKKEYPGRQYAVVDGGMNDIRPALY
QAKYEAAANRIGEAHDKTVSINGKCESGDMITWDIPEVKEGDLAVFCGAYGY
SMANNYNRIPRPVAVFVENCEAHLVVKRETYEDIVLDELPEFKYGVKQ"
1610. 1616

/gene="lysA"
/citation=[1]
1751. 1757
/gene="lysA"
/citation=[1]
2496. 2502
/gene="lysA"
/citation=[1]
2504. 2510
/gene="lysA"
/citation=[1]
2666. 2694
/note="ORFX19"
complement(2693. 3565)
/note="ORFX19"
/codon_start=1
/transl_table=11
/protein_id="AA67474.1"
/db_xref="GI:410118"
/translation="MKKIMIGWLAALVLLWPKVSLADAANGVYVTLGADLSFSDK
OKVLEDMVPMNATVYVTKNKEHEHYLGKYSISNAIGSRATSSSITAKGSCGNT
THNIGITDEMILNALMTAGVDAAVYVTAPEVSGTALGILNAVEVSSDEAISED
VKQVANOELVYTSLEGDKRIGENNAALIAKIEFFAKNGVDPNKAIDIEKQVDDAASDL
NVTLDQSKNOQLVSLFNNKMKNADIDMGQVSDQDKAKDKITFKIESDEGKNFIQKVID
FVFSIMNAIVSIFK"
3347. 3353
/citation=[1]
3413. 3419
/citation=[1]
3428. 3434
/citation=[1]
3503. 3509
/citation=[1]
3819. 4290
/gene="ppIB"
3819. 4250
/gene="ppIB"
/codon_start=1
/transl_table=11
/product="peptidyl-prolyl isomerase"
/protein_id="AA67475.1"
/db_xref="GI:410119"
/translation="MKTGFFLEDSGNKIEFELYPPAAGTVAEFKLANEGFYDLTF
HVRIGFVSGGCPHGTGCGPYTIKCTEGNPTTHAGALSMNAHAKDKDGGQSFPI
VHEPQHANGVTVGKVTSGLEFAKNNNGDVMEVVEG"
3982. 3988
/gene="ppIB"
/citation=[1]
4261. 4290
/gene="ppIB"
4287. 4320
/note="inverted repeat"
5264. 5467
/note="ORFX1"
/codon_start=1
/transl_table=11
/protein_id="AA67476.1"
/db_xref="GI:410120"
/translation="MQLVGLLEQYGAVALFIVLMLELLALLGGEFGAVLVFDEKLN
WVLSMLAAGLGSGSEHYVLDLV"
5445. 5834
/note="ORFX2"
/codon_start=1
/transl_table=11
/protein_id="AA67477.1"
/db_xref="GI:410121"
/translation="MSWIGYRLGPAALOKAFASRGVLLAVIVLTFEFTVREGDGRKI
ELSFPLFVIGELVEEGIRLRFSLHVAHPLESHVLYVPPSBQTLAAFIYFGCAV
LFRSHKTRLOTVAFFCLLSYSSG"
6022. 6027
/gene="sps"
/note="putative"

conflict

-35_signal

```

gene          6022.. 7323
/gene="sips"
-10_signal    6045.. 6051
/gene="sips"
-35_signal    6683.. 6688
/note="putative"
/gene="sips"
/note="putative"
-10_signal    6705.. 6710
/gene="sips"
/note="putative"
CDS
/transl_table=11
/product="signal peptidase"
/protein_id="AA67478.1"
/db_xref="GI:410122"
/translation="MKSENVSKKKSLLEMAKAIYIAVLALLINFIAPYVDGDSM
YPLAHNERVFNMTVKYIGDFRDIDYLDVGVYVKRIGIGDTEVEMKNDLTI
NGKRVDEPYLAANKRAKODGFHDLTDFGVYKVPDKNYFVMDNRNRSMDSRGLLT
FTKROIAGTSKFVFPENEMRKIN"
terminator    7299.. 7323
/gene="sips"
-35_signal    7388.. 7393
/note="putative; OREX4"
-10_signal    7410.. 7416
/note="putative; OREX4"
CDS
/note="OREX4"
/codon_start=1
/transl_table=11
/protein_id="AA67479.1"
/db_xref="GI:410123"
/translation="MGRIKTKITLLVLLLLAGCYWINDIEIKDVPALIGTSLSE
EEEYTDIEYKVTKIDSEYHGVANGKRIIFNGKRLNODLSDIEKGDKIKAYFESKR
IDGLIKAVKMD"
terminator    7872.. 7911
/note="OREX4"
-35_signal    7937.. 7942
-10_signal    7960.. 7965
protein_bind  8126.. 8137
/note="putative"

Query Match      11.7%; Score 120.4; DB 1: Length 28206;
Best Local Similarity 51.0%; Pred. No. 6.3e-13;
Matches 317; Conservative 0; Mismatches 296; Indels 9; Gaps 1;

```

```

Db      8641 CTGGCAGCCAGCGGAGAGAGCGCTGAATGAATAAATTCTGCACITTTATGAGACAGCGCTT 8700
Qy      412 CCACAATTACAGTGAAGAATGATCTGCAAGTTTAGATGCTAAACAGCGAATGATTAAGA 471
Db      8701 CCGTACGTCAGCTTAAAGCGCGCTGCACGCTTGACGGCAAGATGATACACAGCGGT 8760
Qy      472 CAAAGTCATGATTTACTTAAACAAGAGTGAATGAATGATGCTTAAGTTAAGACATGCA 531
Db      8761 GACAGCAATGATGATCAGCTCAGAGGCTGCAAGACAGATGCTCAGCAATACAGAAACAA 8820
Qy      532 CACGACGCGAGTGTACTGGAAGAAGTACAGTTGAATTAAGATCCACATATACTACA 591
Db      8821 CACCAAGACATTTTAGTCGAGGTGGCAGATGAAGCCGACATCCGAGCTTAACCTGC 8880
Qy      592 CGATTTCAGATGCAAAAACC 613
Db      8881 AGACTGCCGGAATGTAAACAAAC 8902

RESULT      8
BSUB0013/c  218470 bp  DNA  BCT  26-NOV-1997
LOCUS      Bacillus subtilis complete genome (section 13 of 21): from 2395261
DEFINITION to 2613730.
ACCESSION  299116 AL009126
VERSION    299116.1 GI:2634723
KEYWORDS
SOURCE
ORGANISM   Bacillus subtilis.
            Bacillus subtilis
            Bacteria; Firmicutes; Bacillus/Clostridium group;
            Bacillus/staphylococcus group; Bacillus.
REFERENCE  1 (bases 1 to 218470)
            Kunz,F., Ogasawara,N., Moszer,I., Albertini,A.M., Alloni,G.,
            Aaevedo,V., Bertero,M.G., Bessières,P., Boločin,A., Borchert,S.,
            Borris,R., Boursier,L., Briens,A., Brum,N., Brignell,S.C.,
            Bron,S., Brouillet,S., Brusch,C.V., Caldwell,B., Capuano,V.,
            Carter,N.M., Choi,S.K., Codani,J.J., Conerton,I.F., Cummings,N.J.,
            Daniel,R.A., Denizot,F., Devine,K.M., Dusterhoft,A., Ehrlich,S.D.,
            Emerson,P.T., Entlan,K.D., Errington,J., Fabret,C., Ferreri,E.,
            Folger,D., Fritz,C., Fujita,M., Fujita,Y., Fuma,S., Galizli,A.,
            Galleron,N., Ghim,S.Y., Glaeser,P., Goffeau,A., Goldightly,E.J.,
            Grandi,G., Guisepi,G., Guy,B.J., Haga,K., Halech,J., Harwood,C.R.,
            Henaut,A., Hilbert,H., Holsappel,S., Hosono,S., Hullo,M.F.,
            Itaya,M., Jones,L., Joris,B., Karamata,D., Kasahara,Y.,
            Klaere-Blanchard,M., Klein,C., Kobayashi,Y., Koetter,P.,
            Konigstein,G., Krogh,S., Kumano,M., Kurita,K., Lapius,A.,
            Lardinois,S., Lauber,J., Lazarevic,V., Lee,S.M., Levine,A., Liu,H.,
            Masuda,S., Mauei,C., Medigue,C., Medina,N., Mellado,R.P.,
            Mizuno,M., Moestl,D., Nakai,S., Noback,M., Noone,D., O'Reilly,M.,
            Ogawa,K., Ogiwara,A., Oudega,B., Park,S.H., Paro,V., Pohl,T.M.,
            Portetelle,D., Potvillik,S., Prescott,A.M., Presecan,E., Fujic,P.,
            Purnelle,B., Rapoport,G., Rey,M., Reynolds,S., Rieger,M.,
            Rivolta,C., Roche,E., Roche,B., Rose,M., Sadale,Y., Sato,T.,
            Scanlan,E., Schleich,S., Schroeter,R., Scoffone,F., Sekiguchi,J.,
            Sekowska,A., Seror,S.J., Serron,P., Shin,B.S., Soldo,B.,
            Sorokin,A., Tacconi,E., Takagi,T., Takahashi,H., Takemaru,K.,
            Takeuchi,M., Yamakoshi,A., Tanaka,T., Tepstra,P., Tognoni,A.,
            Tosato,V., Uchiyama,S., Vandenbol,M., Vanlier,F., Vassiroli,A.,
            Viari,A., Wambut,R., Wedler,E., Wedler,H., Weltenecker,T.,
            Winters,P., Wiput,A., Yamamoto,H., Yamane,K., Yasunoto,K., Yata,K.,
            Yoshida,K., Yoshikawa,H.F., Zumbstein,E., Yoshikawa,H. and
            Danchin,A.
            The complete genome sequence of the gram-positive bacterium
            Bacillus subtilis
            Nature 390 (6657), 249-256 (1997)
            98044033
            2 (bases 1 to 218470)
            Kunz,F., Ogasawara,N., Yoshikawa,H. and Danchin,A.
            Direct Submission
            Submitted (18-NOV-1997) I. Moszer, A. Danchin, Institut Pasteur,
            Regulation de l'Expression Genetique, 28 rue du Docteur Roux, 75724
            Paris Cedex 15, FRANCE. E-mail: moszer@pasteur.fr,

```

adanching@pasteur.fr Phone: +33 (0)1 45 68 84 41, Fax: +33 (0)1 45 68 89 48

FEATURES

source

Location/Qualifiers

1. 218470

/organism="Bacillus subtilis"

/strain="168"

/db_xref="taxon:1423"

terminator

complement(7. .18)

gene

/gene="cmk"

complement(7. .696)

CDS

/gene="cmk"

complement(22. .696)

/EC_number="2.7.4.14"

/note="alternate gene name: jofc, ypfC"

/codon_start=1

/transl_table=11

/product="cytidylate kinase"

/protein_id="CAB14221.1"

/db_xref="GI:2634724"

/db_xref="SWISS-PROT:P38493"

/translation="MEKKSIAIDGPAACKSTVAKIVAEKKSYIYIDGAMTRATY

ALQENVLDLDEKELAEELKRTDIELITTDGKVFVNGVDTEALRTDEISNQVSLA

AKHRSVDEEKVKRQOQGEKGVMDGDIGTIVLPAAEVKIFLLASVEERAKRYEE

NVKKGVDVNETLIEELIARDRKLDSEREVSPLRKADEALIDTSLISIQEVADKILEA

VEOKSR"

gene

complement(775. .951)

CDS

/gene="ypfB"

complement(775. .951)

/gene="ypfB"

/function="unknown"

/note="alternate gene name: jofB"

/codon_start=1

/transl_table=11

/protein_id="CAB14222.1"

/db_xref="GI:2634725"

/db_xref="SWISS-PROT:P38492"

/translation="MKTFERLLIKLFIQLIILLGVQFLFHQIHIEPYVSKVIQYEGV

DKMEENRITFFH"

gene

complement(996. .1649)

CDS

/gene="ypfA"

complement(996. .1649)

/gene="ypfA"

/function="unknown"

/note="alternate gene name: jofA"

/codon_start=1

/transl_table=11

/protein_id="CAB14223.1"

/db_xref="GI:2634726"

/db_xref="SWISS-PROT:P38491"

/translation="MTEIGENVLLEVEEEMIKKASKAVSINNELLIVPYDVYTG

RTYLHNDMEVTVFVGKDEVPYFRFISRIKGVKDKIOMCLEPPPEKKRRIORQY

VRIDAVLDVQIOPGNEEIRTLSTNISAGIAYVLADGLSFQGESLRLLIRLPEEHS

TRDLEAVVRRIFNDPSEKSRKMTLEYSEIAAGDODALLQYCIROLNKRRARBE"

complement(1711. .3094)

gene

/gene="ypB"

complement(1711. .1733)

terminator

/gene="ypB"

complement(1742. .3094)

CDS

/gene="ypB"

complement(1742. .3094)

/function="unknown"

/note="alternate gene name: jofB, yzua

similar to hypothetical proteins"

/codon_start=1

/transl_table=11

/protein_id="CAB14224.1"

/db_xref="GI:2634727"

/db_xref="SWISS-PROT:P38490"

/translation="MIRGILIAVIGIIVGTGYGWGYKEHOEKDAVILLHAENNYORAFH

ELTYQVQDLHDKIGTITLANSQKSLSPALIDWKRTISEAHNSYSQLPITLVPNKKEE

LSKIGDSFYKTSVRDLQKPLDKNETSLNKLQOSEDIONELRHVQHLVPSKRLAK

MDVEMALASDEKQSDNTIINSFTVKEKNVAFSTGDLGFSFTSKKEEGFSLKGG

OISEQEAQIAERFAPDDNYSIKVVGSKKTRNDVYSISMKDPDKAVIYIMDITKKGK

gene

complement(3129. .4046)

CDS

/gene="sleB"

complement(3129. .4046)

/gene="sleB"

/note="alternate gene name: ypea"

/codon_start=1

/transl_table=11

/product="spore cortex-lytic enzyme"

/protein_id="CAB14225.1"

/db_xref="GI:2634728"

/db_xref="SWISS-PROT:P50739"

/translation="MKSKGSIMACILLESFTITTTINTETISAFNOVYORCATGDDV

VELQARLQYNGYINGKIDGVYGMGIWAVRNQDQDGLVEVGLGAKTKQTLCKSK

YREYVMEQLNGKNETFHYGKIPLKQTPSKAAQOKAQOAEARQKQKQPK

ANANKQONTPAKARKQDAVAAANMPGFSNNDIRLIQAQVYGEAREPEQCOVLAAY

ILNRNSPLFPNSVAGVIFEPLAFVADGQIYQMDNETAREAVDAIINGMDPSEBAL

YFNPPTATSPWIMGRPQIKRIGHIFCE"

complement(4162. .4841)

gene

/gene="ypdC"

complement(4162. .4182)

/gene="ypdC"

/function="unknown"

/note="similar to hypothetical proteins"

/codon_start=1

/transl_table=11

/protein_id="CAB14226.1"

/db_xref="GI:2634729"

/db_xref="SWISS-PROT:P50738"

/translation="MFATISAGIAPGIALLSFYLYKQDYDNEPVMHVLRSFFLYLV

FPIMTQYVLEKENVGGSFEVFSLSGFLSEKLMKIMISVYPHAFDEHYDGV

GASVSGFATLENILNYLIGHGVHAEPVRLILVPSCHALIGVIMGFYLGAKRSADAKR

VWMLTSLVPSPLHSGSYDFILTALSNMIYVYMLPFVFLMFGGLRKAKKARSVMNQV

gene

complement(4929. .5935)

terminator

/gene="ypdA"

complement(4929. .4953)

CDS

/gene="ypdA"

complement(4961. .5935)

/gene="ypdA"

/function="unknown"

/note="similar to thioredoxin reductase"

/codon_start=1

/transl_table=11

/protein_id="CAB14227.1"

/db_xref="GI:2634730"

/db_xref="SWISS-PROT:P50736"

/translation="MIOGRAIITIGGPGCLSAIHLKQIGIDALVIEKGVNYSIYN

PHQTFESSSEKLEIGDAFITEENRPVRIQALSYREVYKRNIRVNAFEVNRVTK

TONNFTVIEITSKEIYTPYCIATGYVDPHNPNGVGEDLPVFNHFKGHPYEDKD

VWIGKNSVDALDELKSGARVTVLYRGENYSPSIPVILPEFELVANGYIRMEFG

ACVEKITEENVEVRSGEKELITIKNDVFPAMGYHDDHOLEKIGVEIDKETGRGFEN

EEMENNVGCVFIAGVIAANNANETFIENGFRHGHIAELAKKENH"

complement(5995. .6016)

terminator

/gene="ypcA"

complement(5995. .7324)

gene

/gene="ypcA"

complement(6044. .7324)

CDS

/gene="ypcA"

complement(6044. .7324)

/function="unknown"

/note="similar to glutamate dehydrogenase"

/codon_start=1

/transl_table=11

/protein_id="CAB14228.1"

/db_xref="GI:2634731"

/db_xref="SWISS-PROT:P50735"

/translation="MAADRTGHTEDKLDVLKSTQTVYHKLKLEKLGIPDEYELLKE

PKMLTVKIPVMDDSVKIFGTGRAHDSVGTGIRFHPNVTKEKEVKAVALSLIW

gene
CDS
gene
CDS
gene
CDS
gene
CDS
gene
CDS

TPYFKRRISLSVSLGIFLLLEGAKESHDFEFKHSHVIEESPYGKLSHKTPGN"
complement(2134,.3552)
/gene="ag_128"
complement(2134,.3552)
/gene="ag_128"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAC06502.1"
/db_xref="GI:2982882"
/translation="MORLIIVNPNEPYSLVIKKYVKLIPILVNMLYLVENIMSMILYL
SISPSAANGAGFSLIIMLFYSLSMAISGCTNNLIAQGVAKRDPSPILINGLFSF
LISLPFLFYCGDFVLFLKAVLGASFTVNSLAKKEITLPFWNIPPIGELENTFYGGANNA
GDTKPVMVAIIIMNLTHIGTAYTLINGFFGLPKGVBGWGINISELTAAFIYTFLL
IFEKRPPEHRLERPEKLFLKMYRLGTPTALRAITTTSENFVGELFKFGDKVLAARQ
ICLRLESTSEPMGFSGEMASTLAGONGCANRYGMVAIVATSHPAFNALVSLGLGL
ILEPLHYLVPSRDEVEIEMASYLDQIYGISOPMAAAYSISGLAKGKGXKHIPLFN
ISSMFLPIYPSFLKTIYHSPLVPWCMETETAVALFYTVRKVYVKGLDLPFOKE
EPEBEDNACGKPLEDVASKYT"
3567..4031
/gene="ribH"
/note="ag_132"
3567..4031
/gene="ribH"
/codon_start=1
/transl_table=11
/product="riboflavin synthase beta subunit"
/protein_id="AAC06489.1"
/db_xref="GI:2982869"
/translation="MQITYEGKLTAEGLRFGLVASRPNHALYDRIVEGAINDICIYHGGR
EDDTLVPVGSWEIIPVAGELARKEDIDAVIALGVLRGATPHFDITASVSGLKNL
LSLERKRITPFVGYADTLEQAIERAGTRHGNGKMEALSIAEMANFLSR"
4028..4474
/gene="nusb"
/note="ag_133"
4028..4474
/gene="nusb"
/codon_start=1
/transl_table=11
/product="transcription termination Nusb"
/protein_id="AAC06491.1"
/db_xref="GI:2982871"
/translation="WRIRKRGARDPAFLVLYWMDLRGENPGCLFEVNYEKKIKKDAY
EYAKKLVDPAVRHIEIDSIIIEKHLGMSIDRLGYVENALRIQVABLIFLKSPERP
VFVIDVDVKRYADEDAKGFVNGVLSAIYKAVITTSKEEPSLKE"
4676..5131
/gene="ag_134"
4676..5131
/gene="ag_134"
/codon_start=1
/transl_table=11
/product="putative protein"
/protein_id="AAC06500.1"
/db_xref="GI:2982880"
/translation="MRQYSVIGSSKASEEYEPARYLQKELAKNLNVCCGRTGVE
AVCKAKREGGLTIGIMPSTDYGHGANPPYDIKVNTGMMNMNPIPVASSSEWLVAGVGN
YGTLEIYAILALIKGYVVGYKTKHVEGEVESLEMLKRVDEFEGRL"
5128..6075
/gene="nuem"
/note="ag_135"
5128..6075
/gene="nuem"
/codon_start=1
/transl_table=11
/product="NADH dehydrogenase (ubiquinone)"
/protein_id="AAC06490.1"
/db_xref="GI:2982870"
/translation="MKVFI TGATGFGVGRHIVRELLNRNGYEYHAGVRNLSKERLFGNQ
KGVGVNDKSDSIREALGKNPDVILIGILEKKKGITFERVHYGHRKNLNVEFS
KGFNKRLTFSAISTHDAPRSYHOTRKMAEREVINGLWVTIFRPSIILGPQKLF
EDMKITTYIIPVALPDFGNQOFQPDVYRDVACAABALKAPEIDRKIYELCSGRVYT
KRELADIFFSWDRKVLMI PMPKRLMTFAGLVIRKIILEPFPSSDQIKMMKPNVCV
LKDLEINSIGISCRBIPYEESLKMLENFDIIINK"

Query Match	Best Local Similarity	Matches	Score	DB	Length
1	11.3%	386	116.4	1	14364
8205	50.8%	0	3.7e-12	361	13
61	50.8%	0	3.7e-12	361	13
8145	50.8%	0	3.7e-12	361	13
121	50.8%	0	3.7e-12	361	13
8085	50.8%	0	3.7e-12	361	13
131	50.8%	0	3.7e-12	361	13
8035	50.8%	0	3.7e-12	361	13
241	50.8%	0	3.7e-12	361	13
7965	50.8%	0	3.7e-12	361	13
295	50.8%	0	3.7e-12	361	13
7905	50.8%	0	3.7e-12	361	13
333	50.8%	0	3.7e-12	361	13
7845	50.8%	0	3.7e-12	361	13
412	50.8%	0	3.7e-12	361	13
7785	50.8%	0	3.7e-12	361	13

Db 788 GGAGAGGCCATGCTGATGTCGACGCCATCATATGCGACAGCAGCCCTAGCAAAAGCGCT 847
 Oy 181 ACATTTATATTAAGTTAGAGCCATGATGATTTGGTTCAACACCAACCCCTGTTAAC 240
 Db 848 GATTGTATGTAGCGCTTGAGCCCTGACGCCATTACGAAACACCCCATGTGCCGAA 907
 Oy 241 AAAATATTGATTGTAGATAGCAAAAAGTATACGCAACAAAG-----ACATTCG 294
 Db 908 CTGATCATGAATCAGGAGATAAAAAGGCTGTTTATCGCTGCGAAGATCCCATCCGCTT 967
 Oy 295 TTAGACACACATGTGTATGAGACGTTACGGGCTCAGCGTATGAGTTGAAT---GCGTT 351
 Db 968 GTAGCCGGAAGAGCATCACATGCTGGAAGAACGCGCATCGAAGTGAAGAACGCGGCTT 1027
 Oy 352 GATGATGAAGGGCATCACATTTATACCAAGACTTTTAAAGCAAAAGCAACCACTG 411
 Db 1028 CTCAGGCACAGCGAGAGAGCTGAATAAATGTTTCTTCACCTTCAGCGGACCGACTT 1087
 Oy 412 CCACAAATTCACATGAGATATCTGCAAGTTTATGATGTTAAACAGGAAATGATATGA 471
 Db 1088 CCGATGTACCTGAAAGAGCGCGCAAGCCTTACGCGGAAACAGGACGATGAACCGGT 1147
 Oy 472 CAAATCATGATGATTACTTAACAAAGAGTTAAACAGATGTTATATTAAGTAAACATGA 531
 Db 1148 GACAGCAATGATGATTACATCAGAGCGCGCAAGGCTTATGCCAGCATACAGGAATCG 1207
 Oy 532 CAGAGCGCATGTTACTGGAAGACGTACAGTTGATTAATGATGTC 578
 Db 1208 CATCAAGATGTTCTGTGTGGAGCAGCGACAGTCAAAACCGATATATCC 1254

RESULT 11
 LOCUS U32775 18073 bp DNA BCT 29-MAY-1998
 DEFINITION Haemophilus influenzae Rd section 90 of 163 of the complete genome.
 ACCESSION U32775 L42023
 VERSION U32775.1 GI:1573951
 KEYWORDS
 SOURCE Haemophilus influenzae Rd.
 ORGANISM Haemophilus influenzae Rd.
 Bacteria: Proteobacteria; gamma subdivision; Pasteurellaceae; Haemophilus.

REFERENCE 1 (bases 1 to 18073)
 AUTHORS Fleischmann,R.D., Adams,M.D., White,O., Clayton,R.A., Kirkness,E.F., Kerlavage,A.R., Bult,C.J., Tomb,J., Dougherty,B.A., Merrick,J.M., McKenney,K., Sutton,G.G., Fitzhugh,W., Fields,C.A., Gocayne,J.D., Scott,J.D., Shiley,R., Liu,L.I., Glodek,A., Kelley,J.M., Weidman,J.F., Phillips,C.A., Spriggs,T., Hedblom,E., Cotton,M.D., Uterback,T., Hanna,M.C., Nguyen,D.T., Sauder,D.M., Brandon,R.C., Fine,L.D., Fritchman,J.L., Fuhmann,J.L., Geoghagen,N.S., Gnehm,C.L., McDonald,L.A., Small,K.V., Fraser,C.M., Smith,H.O. and Venter,J.C.
 TITLE Whole-genome random sequencing and assembly of Haemophilus influenzae Rd
 JOURNAL Science 269 (5223), 496-512 (1995)
 MEDLINE 95350630
 REFERENCE 2 (bases 1 to 18073)
 AUTHORS Tatusov,R.L., Mushegian,A.R., Bork,P., Brown,N.P., Hayes,W.S., Borodovsky,M., Rudd,K.E. and Koonin,E.V.
 TITLE Metabolism and evolution of Haemophilus influenzae deduced from a whole-genome comparison with Escherichia coli
 JOURNAL Curr. Biol. 6 (3), 279-291 (1996)

REFERENCE 3 (bases 1 to 18073)
 AUTHORS White,O., Clayton,R.A., Kerlavage,A.R. and Fleischmann,R.D.
 TITLE Direct Submission
 JOURNAL Submitted (25-JUL-1995) The Institute for Genomic Research, 9712 Medical Center Dr. Rockville, MD 20850, USA
 REFERENCE 4 (bases 1 to 18073)
 AUTHORS White,O., Clayton,R.A., Kerlavage,A.R. and Fleischmann,R.D.
 TITLE Direct Submission
 JOURNAL Submitted (27-SEP-1997) The Institute for Genomic Research, 9712 Medical Center Dr. Rockville, MD 20850, USA

REMARK The H. influenzae sequence has been updated by R. Fleischmann. New database matches have been assigned, product names have been improved, and a number of frame shifts have been corrected. We gratefully acknowledge the work of Tatusov et. al. We have incorporated their annotation into the /notes fields of the corresponding H. influenzae genes

REFERENCE 5 (bases 1 to 18073)
 AUTHORS White,O., Clayton,R.A., Kerlavage,A.R., Fleischmann,R.D., Peterson,J., Hickey,E., Dodson,R. and Gwinn,M.
 TITLE Direct Submission
 JOURNAL Submitted (28-MAY-1998) The Institute for Genomic Research, 9712 Medical Center Dr. Rockville, MD 20850, USA
 REMARK The whole genome was shifted by 588 nucleotides for a new start on Sep 30, 1996 this sequence version replaced gi:1221677.

COMMENT Location/Qualifiers
 FEATURES
 source
 1..18073
 /organism="Haemophilus influenzae Rd"
 /db_xref="taxon:71421"
 complement(146..1327)
 /gene="HI0929"
 complement(146..1327)
 /gene="HI0929"
 complement(146..1327)
 /note="similar to GB:M77129 SP:P24196 PID:146678
 PID:882568 GB:U00096 percent identity: 41.10; identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAC22589.1"
 /db_xref="GI:1573952"
 /translation="MKRVGFPPTPDVYQQLNVGFDDYVNLPSDGSNHYSDNVAYE
 FLAEIDRIEDTNNELSMCIDFADEIKKGDENVHFELOKLETSMRNDPEYLG
 RDRGVDGNNIKMFEYNADTPSLLEPAVYQWOMEJOEGKHKRDPKIHLEIKHF
 QFLKQSGKTDPHLSAMODAGREDNGNDYLDVAVYNGMNIHQALVEDIGNSTK
 FVDLNDQPLEMLFKYPLEWSHAFARHTTAEFTLEPAKMLLSNKAALAKLMAR
 YNHPHLPAYTPPELPRLDSMWVKPLGREGANVFEKNNGVFAAKGSEHSTF
 YGNSVYVYQKKEFLPSFDGMYPIIGSWVGVACMGRLREDFATVAGDSHPHPYFV
 P"
 complement(1328..1948)
 /gene="HI0930"
 complement(1328..1948)
 /gene="HI0930"
 /note="hypothetical protein; identified by Genemark;
 putative"
 /codon_start=1
 /transl_table=11
 /product="H. influenzae predicted coding region HI0930"
 /protein_id="AAC22600.1"
 /db_xref="GI:1573966"
 /translation="AARRKNHSIKIENNEKENOILVSLIALLGCGSEQVORDVQ
 SIDDCLAMKKELEADKNTESTOKTATTOOGLINRDGNALSAVKNPAAENNAQ
 AKQSERRAESTTKAESTDSLAGATAGVIGTMAKARLISFEGPEYHGNRAVTTPTG
 QVVOPTNRSVGRPMLVGNAGSMNSKPVSRGFSPPKTRHSSGG"
 complement(1921..2406)
 /gene="HI0931"
 complement(1921..2406)
 /gene="HI0931"
 /note="hypothetical protein; identified by Genemark;
 putative"
 /codon_start=1
 /transl_table=11
 /product="H. influenzae predicted coding region HI0931"
 /protein_id="AAC22601.1"
 /db_xref="GI:1573967"
 /translation="WKSRTLYKSAVFLYLIFKGRMLHLEDFELGTPKQVGTSH
 TVFEHFAVMEDEGTGYFALDMRONOPIDMLHVNVDSTSNHBEARKLEICWDE
 SGYVALLLINGPFAVDFEARLVGNSKRYPOPDLMSWMTBEITNKQAEQWLGVKTI
 R"
 complement(2424..3734)
 /gene="HI0932"
 complement(2424..3734)
 /gene="HI0932"
 /note="similar to GB:M12843 SP:P08324 GB:X82400 PID:147479

ORGANISM	Vibrio cholerae
REFERENCE	Bacteria; Proteobacteria; gamma subdvision: Vibrionaceae; Vibrio.
REFERENCE	1 (bases 1 to 11695)
AUTHORS	Heidelberg,J.F., Eisen,J.A., Nelson,W.C., Clayton,R.A., Gwinn,M.L., Dodson,R.J., Haft,D.H., Hickey,E.K., Peterson,J.D., Umayam,L.A., Gill,S.R., Nelson,K.E., Read,T.D., Tettelin,H., Richardson,D., Ermolenko,M.D., Vamathevan,J., Bress,S., Qin,H., Dragoi,I., Sellers,P., McDonald,L., Utterback,T., Fleischmann,R.D., Nierman,W.C., White,O., Salzberg,S.L., Smith,H.O., Colwell,R.R., Mekalanos,J.J., Venter,J.C. and Fraser,C.M.
TITLE	DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae
JOURNAL	Nature 406, 477-483 (2000)
REFERENCE	2 (bases 1 to 11695)
AUTHORS	Heidelberg,J.F., Eisen,J.A., Nelson,W.C., Clayton,R.A., Gwinn,M.L., Dodson,R.J., Haft,D.H., Hickey,E.K., Peterson,J.D., Umayam,L.A., Gill,S.R., Nelson,K.E., Read,T.D., Tettelin,H., Richardson,D., Ermolenko,M.D., Vamathevan,J., Bress,S., Qin,H., Dragoi,I., Sellers,P., McDonald,L., Utterback,T., Fleischmann,R.D., Nierman,W.C., White,O., Salzberg,S.L., Smith,H.O., Colwell,R.R., Mekalanos,J.J., Venter,J.C. and Fraser,C.M.
TITLE	Direct Submission
JOURNAL	Submitted (14-JUN-2000) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
FEATURES	Location/Qualifiers
SOURCE	1. 11695
	/organism="Vibrio cholerae"
	/strain="N16951"
	/serotype="O1"
	/db_xref="taxon:666"
	/chromosome="1"
	/note="biotype: El Tor"
	/complement(97.825)
	/gene="VC2260"
	/complement(97.825)
	/gene="VC2260"
gene	/note="Similar to GB:ID26562 SP:P03351 GB:V00343 PID:42842 PID:473824; identified by sequence similarity; putative"
CDS	1..start=1
	/codon_start=1
	/transl_table=1
	/product="ribosomal protein S2"
	/protein_id="AAP5404.1"
	/db_xref="GI:9636822"
	/translation="MASYSMRDLTAGVHFGHQTRYNPKKQIFGARNRVHIINLEKTVPMFEALAEKLVGEKKGVLFVGTGRRASSVKEAALASQYVYNNWGLMGTTNWKVPSIRKLKLEVOSTGDTGFDKLTTRKALRTREMEKLEKSLGIDKMGDLPAPALFVIDADHEHAIRKEANNLGIPIPAVVDVNSPGCVYIIPGNDAIARAQVLYNMAAQAINEGRKDYAAVAERKDGEVAE"
gene	1166..2008
	/gene="VC2261"
CDS	1166..2008
	/note="VC2261"
	/note="Similar to GB:ID26562 SP:P07906 GB:M15106 PID:146727 PID:473823; identified by sequence similarity; putative"
	/codon_start=1
	/transl_table=1
	/product="methionine aminopeptidase"
	/protein_id="AAP5405.1"
	/db_xref="GI:9656823"
	/translation="MSIKKVAVEIEKKRAVAGRLAAEYLEKIEPHVACGVTTELDQICHKITVEOGAIPAPLNVHGFPGKICSTINIHVCHGIPASDITFGQIQRPVLRDDQILNLDITVKNQYGDTSKMFLLDGVDSEKRLCHVAQOECLYLAKQVKPVGQGEJGTTIKKIDITKNNKNNRPKRSIVRDYCGGIGIAEHPHPOVYLYNRSRTVLRCEMITTIEPKNVGKFCRCRDEDSWTVYIADCKSKASQWMTHTLVATGCEITLTSSESLPRLI LNN"
gene	2098..4743
	/gene="VC2262"
CDS	2098..4743
	/note="VC2262"
	/note="Similar to GB:M96431 SP:P27249 PID:146155 PID:43933 PID:473822; identified by sequence similarity; putative"
	/codon_start=1
	/transl_table=1

/product="protein-p-II uridylyltransferase"
 /protein_id="AAF95406.1"
 /db_xref="GI:9656824"
 /translation="MHGRAMPYQSPITFOEPDLTVESLKQOLESFTEYKOEFDHNP
 VTDVIGSEYMDLLHRLMOFGFDELVEYSLVAVGSGELPLSDIDLVLSSQ
 PLSEOVAKIISQFLTLDLKEIGHAVRTVGOCAEIGKADLTATNLQEAIRLGG
 ETFHRLKVIISSESPSEITFYQAKVPROKRRHARYHDTTYNLEPDIKSTPGGRDLH
 TLSVARHFEATSLVEMSRFGFLTDAYRELVLCODELNVFRALHLEKRYNDRIL
 FAHQVARNHGFEGEENRGJEMMKKEFFRLRVAVLENKMLTIPKALINNEEAK
 AVIIDDDQSRGNMIEARKPALFOAREPTIIDMVELHMAVDSTIESVAPATROQRTAR
 FLYNVEHTLSIARAREKEITELVRNPALEKASQNHKIGVLAAYIPONNOIYGMOFPI
 RRLVVDLHTRILKRIHLEFSDANNHDHPICCEIYEPKELLILAAIFHDIGSKR
 GGDHEIETADAFDPCITFHGSKSEPAKIVANLYNNHILMSTAAORROIDPYDIIEER
 KVRREEMETLVCITVADICATPELNMESKRRTLDLEAFISTORALRGLENVDR
 ERIRINQOMASALLKKEGFSRELEVMQRYKAYFLRHTKQIAMCTHLLRHEDSS
 KPLVLSKATAREGTEVEIYTRQDAALFATVADLRNLNVHDAQIMASQDQVLEPT
 FMVLQNOQALIEEDRHQALIRHLVHVEDGPTTQKARIIRNLQHFVKQVDFLT
 KSKRTLMEFVALDPPGLATVAGATFELNLDLAAKITITIGEAEDFLITLNQGR
 LNEEBOHLREKLIHVAELAPSAO"
 complement(4818. .4916)
 /gene="VC2263"
 complement(4838. .4916)
 /gene="VC2263"
 /note="identified by Glimmer2; putative"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="AAF95407.1"
 /db_xref="GI:9656825"
 /translation="MSMLMGTHSLAAVLODOIWVYVFTKMAULL"
 4989.5402
 /gene="VC2264"
 /gene="VC2264"
 4989.5402
 /gene="VC2264"
 /note="identified by Glimmer2; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAF95408.1"
 /db_xref="GI:9656826"
 /translation="MRVILTKRLPMFPHLVGLINEPTQIERYSROBAHKDVLYKE
 KKKQKELPAKSVKKEYPQIKNVLDSSHQYKEVTELRNLITLIDELKNTIPPEVM
 GEVDKQKLTLDLHLEKVAKTAIEADLQKLN"
 complement(5443. .5952)
 /gene="VC2265"
 complement(5443. .5952)
 /gene="VC2265"
 /note="similar to GB:U00096 PID:1773102 PID:1786620
 PID:2329841; identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="phosphatidyglycerophosphatase A"
 /protein_id="AAF95409.1"
 /db_xref="GI:9656827"
 /translation="MSDPRNSLUSNPWHLLATFGGSLGYPTGTLASVPEYLM
 LAQVLIYAAVVAVFASLDGIKICQVSDMOVHDHSIYVDEPAGWIMLIVPLQV
 LPEVPMKVLVGFVLFREFDMLKWPMSIWDLRKVHGFGIIMDLVAGVMAAISLAAI
 GYVLEMLPV"
 complement(5949. .6953)
 /gene="VC2266"
 complement(5949. .6953)
 /gene="VC2266"
 /note="similar to GB:U00096 SP:p77785 PID:1773101
 PID:1786619 PID:2329840; identified by sequence
 similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="chiamin-monophosphate kinase"
 /protein_id="AAF95410.1"
 /db_xref="GI:9656828"
 /translation="MQRENOQVMMHGEFNLIDKYFNSRQAKRDVHLALGDCAIYK
 VEPNSRVAISDITVAGTHFLAQNPMVAHKAALASNISDLAAGATPAWVSLTLP
 BIDAAMVADDAFFELAKYVNDLIGDGTTKFLSTLYVQGLPGEQAMLRGAVG


```
/gene="TM1816"  
/note="similar to GB:Pyro.h percent identity: 65.04;  
identified by sequence similarity; putative"  
/codon_start=1  
/transl_table=11  
/product="conserved hypothetical protein"  
/protein_id="AAD36879.1"  
/db_xref="GI:4982398"  
/translation="MIIAPVSENGKSGSPISEHGRAPYAFYVKNNAIADISVEE  
NPAGDDHVGAVPNFVKERGAELVIRIGRAIAAFEMGVKIVKASGVIEEYVNO  
YLSGQDKDSYEVHDHHEHH"  
/complement(1011..3608)  
/gene="TM1817"  
/complement(1011..3608)  
/gene="TM1817"  
/note="similar to GB:M16318 SP:P11931 PID:143797 percent  
identity: 72.43; identified by sequence similarity;  
putative"  
/codon_start=1  
/transl_table=11  
/product="valyl-tRNA synthetase"  
/protein_id="AAD36880.1"  
/db_xref="GI:4982399"  
/translation="MAELSTRYNPAEIEFKWRYWEKGYFTPKGCEKFSIYPPN  
ITGRIMGHALNITLQDIYVRKRMKGYDLVAVPEDHAGIATONAVEKFLIQTGKT  
REIEGREKEFLIETWEMANKYREIREIOIKALGASVDMTRERPTLDEGLSRAVRKVE  
LYRGLIRGKITYIVWCPRCKTVLSDEVEYKHKSKLYKKYKPKDSDEIYVATTA  
PEMIGDVAVAHPEDERTKNFVGKTLILPLGREIPIVADYVDPKGTGAVKATPA  
HDPNDYLAQRNLNPMIEIFDNNRINNGKYGKGLDYERAKELVKDLSOGFLVKI  
EDYHSVGHCRDVTIEPKLSDQWFSSTKPLAKRAIYAVNGEIRFPPEMTKLYLN  
MWYELRDMCISQLMWGRIPWYQCDCGHLNASEDYENKCGSTNLKODEVDLDT  
MSSALPFCSTLGMPEETEDLKRYPTDLVVGPIITFFWVARMIMGMGFEMNDKPS  
HVIYIQLVRDKYGRKMSKSLGNDIPLEVIDEYGDMPRFTLALIAAGRIKIDPRY  
DFAIKFRANKINATRYVLMNDEIKVEYLEMLKTVDMWILFRKITYEETNALEN  
FENIAARTIYNEFMDPCDWYIEASKPRLKTEERNLVQTVLVKVLDAASLRLLHPMP  
LTELMOQLPVAGEEIIIAKWPETEIRELIDEAEKEFTRLMNAVGVNVAENMLPQ  
SORVYVYIKGEVTEEEELLTKGNIIEEVSNEKPKPTATAYEEIEIYAVDGLD  
IDFEKERELQIMKEIKQELDRLEKLANKDVEKAEVEVEETKEKLNRRERLAR  
LESTIRDE"  
/complement(3640..4140)  
/gene="TM1818"  
/complement(3640..4140)  
/gene="TM1818"  
/note="similar to GB:AE000783 percent identity: 61.07;  
identified by sequence similarity; putative"  
/codon_start=1  
/transl_table=11  
/product="chromate transport protein, putative"  
/protein_id="AAD36881.1"  
/db_xref="GI:4982400"  
/translation="MLKLAIFPKIGFISFGGNAYIVGILKNELVTGGSFSPESOA  
VSIADMPGPVAINLATYTGKFFGLIGAVLNTLFLDAPLIVITTAIFLKKYKLR  
VRLKALGATTTLLIVTLSSLSSVQNPVILLSSAAVFCSEFKVHPLFIIFGCVI  
GAILGF"  
/complement(4144..4716)  
/gene="TM1819"  
/complement(4144..4716)  
/gene="TM1819"  
/note="similar to PID:687689 percent identity: 55.06;  
identified by sequence similarity; putative"  
/codon_start=1  
/transl_table=11  
/product="chromate transport protein, putative"  
/protein_id="AAD36882.1"  
/db_xref="GI:4982401"  
/translation="MELICEGSKSRIEVEGSSNLIGRLFFLRISALTIGGAYAMIP  
VMKVELSGLTEKEFRIMSTAOVPGPIAFNTAVVGRRLAGIYGAISGLAVVL  
PPFAIVAVAEVIRLGSISYVSRIRGAVASIIIGVGSVLYRLVRNQRMLYRAIMI  
GVAVFVLLNSLVIPVILVLLILEV"  
/db_xref="GI:4982402"  
/gene="TM1820"  
/db_xref="GI:4982402"  
/gene="TM1820"
```

```
/note="similar to GB:AE000657 percent identity: 76.48;  
identified by sequence similarity; putative"  
/codon_start=1  
/transl_table=11  
/product="GMP synthase"  
/protein_id="AAD36883.1"  
/db_xref="GI:4982402"  
/translation="MVLVVDYGSQYSRLTRIRENENYSEFVPPDDKVDLSKYDAVI  
LSGSPRSVYEDDAKLLPEMFOEYKGPVLAICYGQQLVKELGSGVRGRREGYGLVE  
LSRPIETGEYREKHYVMSHODEVYRLPEGEHPIAVSETGVIATATGKRRLQIHP  
EHHTEYGDPMISNPLFNCKLEKNMIGDLVEKIRHIEETIGNKRAIILASGVVS  
SVAALVHRAIGKMLVCVFDHGLLRNRENEREVRKHEHDMNVVDVAKRRLLEL  
RGVDPPEKRRKIIIEEFIRVEEAKKHDEYVQGTIYSDVIESASGATTAKIKSH  
HNVGDLPEKMKLKIPEPLDLDFEVRKGVKGYIGIPRIINRHPFPGGLAVRYLGEV  
TEELLELRADYFIETILRKHDYVDKVMGAFNLLPIKISVYGADARAYEVVALVA  
VNSYGMTADSRIPHDILDAARRITREYKGVGVYDITSRKPATIEWE"  
/db_xref="GI:4982403"  
/gene="TM1821"  
/db_xref="GI:4982403"  
/gene="TM1821"  
/note="similar to GB:L77117 SP:Q58134 PID:1499541 percent  
identity: 72.33; identified by sequence similarity;  
putative"  
/codon_start=1  
/transl_table=11  
/product="conserved hypothetical protein"  
/protein_id="AAD36884.1"  
/db_xref="GI:4982403"  
/translation="MELLKELEERIREANPLVEORFEFRKLGEECTEEDLFCGLSE  
CVLVAANSABEGIAQAQKEIGGFYHLPLEELAEKLRVGHRYPOKRAEFTVENKRLG  
KLKLVGSGDPSREFPLVRNAKGIGMKEASHFTLNTGVEDLAIIDKRVILMRKHGLI  
QELPKGSKRYLVEELRKVAAPFESGKFDLYLMYLVKQVDR"  
/db_xref="GI:4982404"  
/gene="TM1822"  
/db_xref="GI:4982404"  
/gene="TM1822"  
/note="similar to GB:U09005 SP:P40605 PID:507724 percent  
identity: 58.45; identified by sequence similarity;  
putative"  
/codon_start=1  
/transl_table=11  
/product="f1sh protease activity modulator HflK"  
/protein_id="AAD36885.1"  
/db_xref="GI:4982404"  
/translation="MRKYVIVYVIVIGYFLNGYVGVSEVALKTFGRFYSVYS  
GIHYHLPYIQSHTVDVITVRKIEIFRSIQGERISTOSPOEAIMITGDNLVYS  
EAVQYVRKDPVAFAFNITEADSLVRFTESSVAREKAMNSIDVLTSGDEIGFKTA  
QMLQEIIDSYNCIGIVENVYLOEVVPPVDVADFDVNNAROKERLINSEARXYANDY  
VPKAGOGAEOELIRQAEVAYAOEVYKALGEMKRFEEVLEESKADITRRKMLDALQS  
LLEKSENKVFVFGDSLNTLNTSLDLKMGK"  
/db_xref="GI:4982405"  
/gene="TM1823"  
/db_xref="GI:4982405"  
/gene="TM1823"  
/note="similar to GB:U09005 SP:P40606 PID:507725 percent  
identity: 61.97; identified by sequence similarity;  
putative"  
/codon_start=1  
/transl_table=11  
/product="f1sh protease activity modulator HflC"  
/protein_id="AAD36886.1"  
/db_xref="GI:4982405"  
/translation="MKIWMISLLIILVVGAILLFSFFYVLDOTOQAVVRFKIVAY  
ETEPGLHFKQPVNDVVRFDKRLILYDIEPKIIIAKKTLLVDTYLMKIKAEAI  
KSLKSVKLAIDRIDENVYSHVRNTEFAANDEIIESEKDLREVTALSEEDKDFQI  
EVDVVRVKAHLDAENENKAYVERKARERYSIAQIRAECEKARKIRAEADKTKAVLI  
AAEOSKQEOIKGCGEASAVKIYAEVFSKDDFVEFWMTMEVYSIEKGIILIDGDELDA  
LKYIKTK"  
/complement(8704..9519)  
/gene="TM1824"  
/complement(8704..9519)  
/gene="TM1824"  
/note="similar to percent identity: 0.00; identified by
```


sequence similarity: putative"
/codon_start=1
/transl_table=11

Query Match 8.18; Score 83.4; DE 1; Length 12811;
Best Local Similarity 50.28; Pred. No. 3.9e-06;
Matches 322; Conservative 0; Mismatches 301; Indels 18; Gaps 4;

QY 1 ATGATATATGCGATTCACAAATATGTTACAGTCAACAGGCTTAATCCACC 60
DB 12711 ATGAAAGAGCGATAGAGCTTGCAAGAAAGACTTGGAGAGGTAATCCACACC 12852
QY 61 GTTGCGCCCTGTTAGTAAAGAGTATGATGTTGTTGTCACACTTGAGAAA 120
DB 12651 GTTGCGCGGCTGTTGAAAGAGCGAGATATCCGGAGGTTTTCATCCCTATTT 12592
QY 121 GGTGACAAAGCATGCGGAGGTTCAAGCACTTGATATGCGACAAATAATGCTGA----- 173
DB 12591 GGAAGTCCGACGCTGAAGAGATGCAATAGAAAGTGCAGAAAGAAAGAGATTTA 12532
QY 174 -AGTGGCGAGATTTATATACCTTAGAGCCATGATGCTATTTGGTTCACACCAACC 231
DB 12531 CGAGGTCAACTCTCATAGTATACCTTGACCTGATCATGGAAGACACCTCCA 12472
QY 232 TGTGTTAACAATAATTATGATTTAGATAGCAAAAGTATATACGACAAAGA---C 288
DB 12471 TGCAGGATGTCATCATAGAAAGTGAATAAACCCTGATGATGGAACAGGATCCA 12412
QY 289 AATTCGTTAGACACATGCTGATGAGACGTTACGGCTCAGCTATTTAGCT---TGAA 345
DB 12411 AATCCTGTTTCGGGAAACGGGTGGAATAATTCAGAAATCAGCTATTTAGATAGAA 12352
QY 346 TGGCTTATGATGAGACGGGATCATCAATTTATACCAAGCTTTTAAAGCAAAAGCAAG 405
DB 12351 GGTGTTTGGAGAGAGAGGTAAGAGCTGTGAGTTTTCACACTTATGTGACAAAG 12292
QY 406 CACTGCGCAAAATATACAGTGAAGTATCTGCAAGTTTATGATGTTAAACAACGATGAT 465
DB 12291 AAAAGCGCTTTCGTTGGGCTGTAAGTACGCTCCACCTTAAATGGAATAATGATCATC 12232
QY 466 AATGACAAAGTCAATGATTTACTACAAAGAGTTTAAACAAGTGTCTATTAAGTATAGA 525
DB 12231 AGAGGATATTCAGATGATGATCAGACAGAAAC---TCAGATTCAGAGTTTCAAGATGAGA 12175
QY 526 CATGACACAGCAGCAGTGTATCTAGCAAGAGTACAGTGTATTAATGATGATCCCAATAT 585
DB 12174 AATCATCTACTCTGCTGTTCTGCTGGTCTGGAACGGTTTGAAGGACAAATCTCAGCTA 12115
QY 586 ACTACAGTATTCAGATGGAATAAACCCCTATAAGTAAT 626
DB 12114 ACCGACAGACTGAAGAAGGCAAGAAATCTGTGAGGCTGAT 12074

RESULT 14
TNE9832 5451 bp DNA BCT 11-FEB-2000
LOCUS Thermotoga neapolitana malg, malg, agla and ribd genes, partial.
DEFINITION
ACCESSION AJ009832.1 GI:3378521
VERSION
KEYWORDS
glu gene: alpha-glucosidase; cyclomaltoextrin glucanotransferase;
malg gene: malg gene; putative; ribd gene; riboflavin-specific
deaminase; sugar transport protein; transmembrane.
SOURCE
ORGANISM Thermotoga neapolitana
Bacteria; Thermotogales; Thermotoga.
REFERENCE
AUTHORS Bererina,O.V., Zverlov,V.V., Lunina,N.A., Chekanovskaya,L.A.,
Dublina,E.N., Liebl,W. and Velikodvorskaya,G.A.
TITLE Gene and properties of thermostable 4-alpha-glucanotransferase of
Thermotoga neapolitana
JOURNAL Mol. Biol. 33, 801-806 (1999)
REFERENCE
AUTHORS Zverlov,V.V.

TITLE Direct Submission
JOURNAL Submitted (29-JUL-1998) Zverlov V.V., Russian Academy of Science,
Institute of Molecular Genetics, Kurchatov Sq., 123182 Moscow,
RUSSIA

FEATURES
source Location/Qualifiers
1..5451
/organism="Thermotoga neapolitana"
/strain="2706-MC24"
/db_xref="taxon:2337"
1..1499
/gene="malg"
/gene="malg"
<1..1499
/gene="malg"
/codon_start=3
/transl_table=11
/product="putative transmembrane sugar transport protein"
/protein_id="CA08866.1"
/db_xref="GI:3378522"
/db_xref="SPTREMBL:O86958"
/translation="VTSGDEGILLQDGNITKSLYRNFLQSFEEVITPVGVKQMS
NIDSPLOKAKELKDQNELAFINSRLDEGLKSLTTLTMAQSKESRISLARYVDLRV
EVEYENKRRVEDLKSNSAQIQLSLYPSKSPVSLYISQCNDSFOTIRKMS
ELSWIEDTRESRMEGTGNADLILNSRKVLDFKQSYPMILDLSIRGFVSEHL
QMLYDLKKNFQEVLTNTAVASRAAGSLMSVPLKELRSDFKRIDGLYRAQIWEQ
KIRHYFLRWVANSIVAGLVIITAVCALAAYPSRRMFRGROYGIALLLIQMFA
IMWVAIYGLKIGQPLFLGLDGLGGLIFAYLGNIAVMYLIGFYDTIPSSLEA
AMIDGATRFQFTKIVPPLAIPILSVIYLPIFGFENFEVLARIILQDVKNYTAGL
WTFSTGATYTEMGLFTAAALIMTMMVILFLSLQKIYIGGLTKGSVK"

gene
CDS

gene
CDS

gene
CDS

gene


```

CDS
4729..>5451
/gene="ribD"
/codon_start=1
/transl_table=11
/product="riboflavin-specific deaminase"
/protein_id="CAA08869.1"
/db_xref="GI:3378525"
/db_xref="SPTREMBL:O86961"
/translation="MYERPMRAIELAKRGRVNPVGAIVVKEGRISGEFHL
FLVGRPAERVAIESAKRKEDLTGTLVLTLPCHHGHTPPCDLIIESGKRYVLD
MDPNVSGSVEKLKRGIEVEEGLVEEVKLEFFTYVTENPITALKRYASTLD
GKIADAEGNSKMTODLRKHAEMRNVYSAVIVGATVLDKDPRLCRKGGRNPTRV
IDRRGIISGGSYRVFEDNARYI"

BASE COUNT      1662 a      1139 c      1401 g      1249 t

ORIGIN
Query Match      7.7%; Score 79; DB 65; Length 5451;
Best Local Similarity 49.5%; Pred. No. 2.8e-05;
Matches 339; Conservative 0; Mismatches 315; Indels 21; Gaps 4;

OY 1 ATGGATTATCGATTCAACTGCAAAATATGTCMAAGTCAAAAGCTGTTAATCCACC 60
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 4744 ATGAAAAGACCAATAGAACTTCCAAAGAAAGGGCTTGGAAAGTGAATCCCAATCTCC 4803

OY 61 GTGGGGCTGTGTAGTATATGAAAGTAGATGTTGTGTGTCGACACT----- 112
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 4804 GTGGGTGCTGTGTAGTGAAGGAAGATATATCGAAAGGGTTTCAATCCCTATT 4863

OY 113 ---TGAAAAGAGTGCACAGCATGCGAGGTTCAAGCACTTGATATGCGACAACAAAT 168
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 4864 TTGGTGGGCCACACGCCGGAAGGGGTGGCATAGAAAAGCCGAAAGGAAGGGAAGAT 4923

OY 169 GCTGAAGGTCGCGCATTTATATTAGCTTAGACCATGTGTCATTTTGTCAACCA 228
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 4924 CTAACTGGAACCACTTGTGTGTCACCTTGAGCCCTGTATCATCTGGGAAACCCCT 4983

OY 229 CCCTGTATTAACAAATTTATTGATTGATAGCAAGCAATAGATATCGCAACAAAAGA- 287
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 4984 CCTGCACAGACTTGTATCATCGAAAGCGGATTAATAAAGTATGTAATCGAATGACGAT 5043

OY 288 --CAATTCGTTAGACACACATGGTGTATGACGCTTACGGGCTCAGCGTATTGA--GGTT 342
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 5044 CCAACCCCTGTTTCAGGAAGTGGTGTGGAATAAATCAAAAACGGAATAGAAAGTGTC 5103

OY 343 GAATGCGTTGATGATGACCGGCATCACAATTATCCACAGACTTTTAAAGCAAAAGCA 402
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 5104 GAAGGAGTGTGGAAGAAAGGTGAAAAAACTCTGCGAGTTTTCATCAGCTACGTACMA 5163

OY 403 AACCACTGCGCAAAATTTACAGTGAAGATCTGCAAGTTTACATGTAACAACGAAT 462
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 5164 GAAACAGGCGCTTTCATCGTTGAAATACGTTCAACTTTTACGGAAGATAGCCGAC 5223

OY 463 GATAATGCAAAAGTCAATGATTACTTAACAAGAGGTTAACAAGATGCTCTAAGTTA 522
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 5224 GCTGAAGGCAACTCCAAATGATCAC---ACAGAGCTTGACACAACAGCCCATGAGATG 5280

OY 523 AGACATCGACAGCAGCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 562
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 5281 AGAAACGATATCTGCGCGTCTTGTGCGTGCAGAAACCTTCTGAAAGATGATCCAGA 5340

OY 583 TATACATACAGTATTCAGATGGAAGAAAACCCATATAAATATATGTAGTCTGGG 642
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 5341 CTGACCTGACAGATGAAGGTGGAAGAAATCCACAAAGGGTAAATCCATCGATCGAGGGG 5400

OY 643 AATAT 647
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 5401 ATTTT 5405
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

RESULT 15
ECNUB
LOCUS      ECNUB      2949 bp      DNA      BCT      26-JUL-1993

```

```

DEFINITION E.coli genes nusB (ssyB) and 3 orfs.
ACCESSION X64395.1
VERSION X64395.1 GI:42147
KEYWORDS nusB (ssyB) gene.
SOURCE Escherichia coli.
ORGANISM Escherichia coli
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE 1 (bases 1 to 2949)
AUTHORS Ito,K.
TITLE Direct Submission
JOURNAL Submitted (06-FEB-1992) K. Ito, Inst for Virus Research, Kyoto
University, Sakyo-ku, Kyoto 606-01, JAPAN
REFERENCE 2 (bases 1 to 2949)
AUTHORS Taura,T., Ueguchi,C., Shiba,K. and Ito,K.
TITLE Insertional disruption of the nusB (ssyB) gene leads to
cold-sensitive growth of Escherichia coli and suppression of the
secY24 mutation
JOURNAL Mol. Gen. Genet. 234 (3), 429-432 (1992)
MEDLINE 93024316
FEATURES
source
misc_feature
    location/Qualifiers
        1..2949
        /organism="Escherichia coli"
        /strain="K12"
        /db_xref="taxon:562"
        /map="9.5 min"
        330..357
        /note="promoter-like sequence p1"
        333..362
        /note="promoter-like sequence p1'"
        391..840
        /note="orf1"
        /codon_start=1
        /transl_table=11
        /protein_id="CAA45734.1"
        /db_xref="GI:42148"
        /db_xref="SWISS-PROT:P25338"
        /translation="MHCFCEAVDTKVIDSLRVGEGSSVRRROCTVNERFTTEVA
        ELVMPRVVSNDRVREPNEKLRSGMLEKRPSPSDVEAIHHSQLRATGERE
        VPSKIGSLVMEQLKLDKVAIYRASVYRSFEDLKEGEELARLD"
        844..1947
        /note="orf2, homologue to B.subtilis ribG"
        /codon_start=1
        /transl_table=11
        /protein_id="CAA45735.1"
        /db_xref="GI:581147"
        /db_xref="SWISS-PROT:P25339"
        /translation="MODEYMAKALKLAORGRFTTHPNPNVGCYIVKDEIVGEGYHO
        RAGEPHAEVHALRMAGEKAGATAYLTPECSHHGRTPCCDALIAAGAVARVASMOP
        NPNVAGGLVRLDQAGIDVSHGLMSEAEIDLNKGFELKRMRTGPYIOLKIGASLDGR
        TMAEGSEOWITSPARRDYOLRPOSHAILTSSATVADDPALTVRSSELDDEOTOL
        YPOUNLROPRIVIDSONRVTVRHIVQDPETWFAKTOEDSREMPETVRLTLPK
        TLDLGLKLDADPQEFKEIRHVGDPVCLHVG"
        1957..1986
        /note="promoter-like sequence p2"
        2036..2506
        /note="orf3, homologue to B.subtilis ribH"
        /codon_start=1
        /transl_table=11
        /protein_id="CAA45736.1"
        /db_xref="GI:42150"
        /db_xref="SWISS-PROT:P25540"
        /translation="NIIIEANVATPDARVAITTAFFNNFINDSLLEGALDLKRGVOY
        KIDENTIVVPGAYELPLAAGALATGYDVIALGTIVIRGTAFFEVVAGASAGSLA
        HVAODEIPVAFTGESIDQAIERAGCTKAGNKGAEMALTALEMIVLKAIKA"
        2526..2945
        /gene="nusB (ssyB)"
        2526..2945
        /gene="nusB (ssyB)"
        /codon_start=1
        /transl_table=11
        /protein_id="CAA45737.1"

```

```
/db_xref="GI:581148"  
/db_xref="SWISS-PROT:P04381"  
/translation="MKPARRRARECAVOALYSWOLSONDIADVEYOFLEAODYKVD  
VLFRRELIAGVATNTMYLIDGLMKRPLYSLDELELOVEKAVLRIRIALYELSKRSDVPYKV  
AINEAIELAKSFQADSHKPFVNGLIDKRAPIRPNKK"  
BASE COUNT      708 a      739 c      825 g      877 t  
ORIGIN
```

```
Query Match      7.6%; Score 78.6; DB 2; Length 2949;  
Best Local Similarity 48.1%; Pred. No.3.5e-05;  
Matches 290; Conservative 0; Mismatches 304; Indels 9; Gaps 2;  
  
QY      1 ATGATTTATGCGATTCACATTCGCAATATGTGACAAAGTCAAGTCAAGTGTATATCCACC 60  
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  
Db      862 ATGGGCGGGCCCTTAAGCTGGCGGACAGAGACGTTTACACCGCATCCCAACCGAAT 921  
  
QY      61 GTTGCGCTGTGTAGTTAATGAAGTAGATGTGTGTATGTGTGACACTTGAGAAAA 120  
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  
Db      922 GTGGGTGCTGTCATGTGCAAAAGATGGCAAAATGTGGTGAAGGTTACCAACCAAGTGG 981  
  
QY      121 GGTGACAAAGCATCGGAGGTTCAAGCACTTGATATGCGCACAAATAATGCTGAAGTGG 180  
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  
Db      982 GGTGACCAACATCGGAGTACGCGGTGCGTATGCGGGGTGAAGAAAGCCAAAGGTGG 1041  
  
QY      181 ACGATTTATATACCTTAGAGCCATGTAGTCAATTTGTTCAACACCACTGTGTAAAC 240  
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  
Db      1042 ACCGCTATGTCACTCGAACCCTGTAGCCATCATGTGCTAGCCACACCGTGTGAC 1101  
  
QY      241 AAAATTATGATTGTAGATAGCAAAAGTAGTATACGCCAANAAGACAAATTCGTTAGAC 300  
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  
Db      1102 GCACCTATCGCGCGTGGCGGTGGCGGTGGCTGCTGATGCAAGATCCTAACCCGCGAG 1161  
  
QY      301 ACACATGGGATGAGACGTTACGGGCTCACGGTATTGAGTTGA-----ATGCGTTG 352  
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  
Db      1162 GTGCTGGGCGGTGGACTTACCCTGCAACAGGCTGGCATTGACGTACGCCAGCGCCTG 1221  
  
QY      353 ATGATGAACGGGCATCACAAATTATACCAAGACTTTTAAAGCAAAAGCAAAAGCAACTGC 412  
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  
Db      1222 ATGATGAGTGAAGCCGAGCAATGTGAATAAGGCTTTCACAGCGATGGCGCGGCTTT 1281  
  
QY      413 CACAATTTACAGTGAAGATAT-CTGCAAGTTTATGATGGTAAACAAGCGAATGATATGGA 471  
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  
Db      1282 CCTTATATTCAAGTTAAACTTGGCGCATGCGTGTGATGGTGGCACGCGATGGCGAGCGC 1341  
  
QY      472 CAAAGTCATGATGATTAACAAGAGGTTAAACAAGATGCTATAGTTAGTACATGGA 531  
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  
Db      1342 GAAAGCCAGTGGATCACTTCGCCCCAGGCGGCGGTGATGTACACTACTGCGCGGCA 1401  
  
QY      532 CACGACGAGTGTAACTGAAGAGTACAGTTGAATTAGATGATCAATATATACTACA 591  
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  
Db      1402 AGTCATGCCATTTTAAACCAAGCGCGCATGCTGGCGGATGATCCTGCTTAACGGTG 1461  
  
QY      592 CGT 594  
      |||  
Db      1462 CGT 1464  
      |||
```

Search completed: November 19, 2000, 04:23:24
Job time: 12787 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 19, 2000, 03:12:14 : Search time 118.16 Seconds
(without alignments)
3271.470 Million cell updates/sec

Title: US-08-978-456-1

Perfect score: 1029

Sequence: 1 ATGATTTGCGATTCAACT.....TACTTACGAAGAGTGA 1029

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 480022 seqs, 187831343 residues

Total number of hits satisfying chosen parameters: 960044

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

N_Geneseq_36.*
1: /SID56/gcgdata/geneseq/geneseqn/NA1980.DAT.*
2: /SID56/gcgdata/geneseq/geneseqn/NA1981.DAT.*
3: /SID56/gcgdata/geneseq/geneseqn/NA1982.DAT.*
4: /SID56/gcgdata/geneseq/geneseqn/NA1983.DAT.*
5: /SID56/gcgdata/geneseq/geneseqn/NA1984.DAT.*
6: /SID56/gcgdata/geneseq/geneseqn/NA1985.DAT.*
7: /SID56/gcgdata/geneseq/geneseqn/NA1986.DAT.*
8: /SID56/gcgdata/geneseq/geneseqn/NA1987.DAT.*
9: /SID56/gcgdata/geneseq/geneseqn/NA1988.DAT.*
10: /SID56/gcgdata/geneseq/geneseqn/NA1989.DAT.*
11: /SID56/gcgdata/geneseq/geneseqn/NA1990.DAT.*
12: /SID56/gcgdata/geneseq/geneseqn/NA1991.DAT.*
13: /SID56/gcgdata/geneseq/geneseqn/NA1992.DAT.*
14: /SID56/gcgdata/geneseq/geneseqn/NA1993.DAT.*
15: /SID56/gcgdata/geneseq/geneseqn/NA1994.DAT.*
16: /SID56/gcgdata/geneseq/geneseqn/NA1995.DAT.*
17: /SID56/gcgdata/geneseq/geneseqn/NA1996.DAT.*
18: /SID56/gcgdata/geneseq/geneseqn/NA1997.DAT.*
19: /SID56/gcgdata/geneseq/geneseqn/NA1998.DAT.*
20: /SID56/gcgdata/geneseq/geneseqn/NA1999.DAT.*
21: /SID56/gcgdata/geneseq/geneseqn/NA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1029	100.0	1029	20	Pyrimidine deaminase
2	1029	100.0	3336	20	Riboflavin-biosynt
3	1029	100.0	3336	20	S.-aureus-riboflav
4	1029	100.0	3336	20	Riboflavin-biosynt
5	1029	100.0	3336	20	S.-aureus-riboflav
6	1000	97.2	7588	18	Staphylococcus aur
7	352.2	34.2	1269	18	DNA encoding a Rib
8	352.2	34.2	1269	20	Pyrimidine deamina
9	120.4	11.7	3367	12	B.subtilis rib (ri
10	120.4	11.7	5567	20	B.subtilis rib op
11	120.4	11.7	5567	20	B.subtilis rib op
12	120.4	11.7	5567	20	Bacillus subtilis

13	120.4	11.7	5567	21	293752	Rib operon of Bact
14	114.2	11.1	14273	19	V52173	Streptococcus pneu
15	112.6	10.9	1101	20	X77179	S. pneumoniae ribg
16	112.6	10.9	3417	20	X77183	S. pneumoniae ribg
17	112.6	10.9	3417	20	X77574	S. pneumoniae ribg
18	112.6	10.9	3417	20	X79558	S. pneumoniae ribg
19	112.6	10.9	3417	20	V72454	S. pneumoniae ribg
20	77.6	7.5	546	19	Z6259	S. pneumoniae deri
21	76.2	7.4	1230	20	X82297	App ribg gene sequ
22	76.2	7.4	4312	19	V27186	Actinobacillus ple
23	76.2	7.4	4312	20	X82299	Nucleotide sequenc
24	54.4	5.3	1686	16	O87587	DNA encoding Leuco
25	51.2	5.0	3095	11	O03875	Sequence encoding
26	47.8	4.6	9636	15	O67190	P. faiciparum tran
27	47.2	4.6	1826	19	V29477	Orpinomyces cellu
28	47.2	4.6	1826	19	V37413	Orpinomyces cellu
29	46.6	4.5	6124	11	O03568	Sequence encoding
30	46.4	4.5	5589	15	O68195	Corynebacterium r1
31	44	4.3	1864	8	N71405	Sequence of ANS-1
32	43.4	4.2	19124	18	T72882	Plasmodium var-7 g
33	43.4	4.2	19124	21	Z98287	Plasmodium var-7 g
34	43.2	4.2	933	20	X61528	B. burgdorferi ant
35	43.2	4.2	1008	20	X61527	B. burgdorferi ant
36	43.2	4.2	910715	20	X20248	Borrelia burgdorfe
37	43	4.2	1206	21	A27366	Sinorhizobium meli
38	43	4.2	4248	21	A27364	Sinorhizobium meli
39	42.8	4.2	162450	21	Z69657	Retinoblastoma bin
40	42.6	4.1	2418	13	Q27886	P.falciptarum GBP13
41	42	4.1	3399	17	T05868	Chicken Leucocytoz
42	41.8	4.1	24601	20	X13160	Enterococcus faeca
43	41.6	4.0	4766	15	O70102	Malariatal PFEWP3 ep
44	41.2	4.0	1864	15	O78892	Aspergillus nidula
45	41	4.0	58407	19	V21210	Methanococcus jann

ALIGNMENTS

RESULT 1	
X59909	
ID	X59909 standard; DNA; 1029 BP.
AC	X59909;
XX	
DT	04-AUG-1999 (first entry)
DE	Pyrimidine deaminase and pyrimidine reductase (ribg) DNA.
XX	
KW	Pyrimidine deaminase; pyrimidine reductase; ribg; antagonist;
KW	agonist; antimicrobial; antibiotic; Helicobacter pylori infection;
KW	H. pylori-induced cancer; gastrointestinal carcinoma; gastric ulcer;
KW	gastritis; ss.
XX	
OS	Staphylococcus aureus.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	1.1029
FT	/tag= a
FT	/product= "ribg"
XX	
PN	W09926475-A1.
XX	
PD	03-JUN-1999.
XX	
PF	23-NOV-1998; 98MO-US24857.
XX	
PR	25-NOV-1997; 97US-0978456.
XX	
PA	(SMK) SMITHKLINE BEECHAM CORP.
PI	Black MT, Burnham MKR, Fedon JC, Hodgson JE, Knowles DJC,
PI	Lonetto MA, Nicholas RO, Palmer LM, Pratt JM, Reichard RW,
PI	Rosenberg M, Tranl CM, Ward JM, Warren RL,

XX WPI: 1999-347572/29.
DR P-PSDB: Y15913.
XX
PT New Staphylococcus aureus polypeptide and polynucleotide useful in
XX the treatment of gastric ulcer and gastritis
PS
XX Claim 6; Page 5; 48pp; English.
XX
CC The present sequence encodes a Staphylococcus aureus pyrimidine
CC deaminase and pyrimidine reductase (ribg) polypeptide. Staphylococcus
CC aureus ribg and its antagonists are used to treat individuals in
CC need of them. Disease related to expression or activity of ribg can
CC be determined by analysing the nucleic acid sequence encoding ribg
CC or detecting the ribg polypeptide in a sample. ribg can also be used
CC to identify antagonists or agonists. ribg, or its related nucleic acid,
CC also has use as a vaccine to induce an immunological response in an
CC animal. Antimicrobial compounds (e.g. agonists and antagonists of ribg),
CC especially broad-spectrum antibiotics, may be of use in the treatment
CC of Helicobacter pylori infection. This should decrease in the advent of
CC H. pylori-induced cancers, such as gastrointestinal carcinoma. The
CC treatment should also cure gastric ulcers and gastritis.
CC
XX
SQ Sequence 1029 BP: 376 A; 162 C; 198 G; 293 T; 0 other:

Query Match 100.0%; Score 1029; DB 20; Length 1029;
Best Local Similarity 100.0%; Pred. No. 3.6e-230;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGATTTATTCGATTCACATTGCAATATGTACAAAGGTCAAAACAGGTGTTATCCACC 60
DB 1 atgattatgcatcattcaacttgcataataatgtaacaagtcacagtgtaattccacc 60
OY 61 GTTGGCGCTGTGTGTTATATAGAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 120
DB 61 gtggcgctgtgtgtatgtaataagtagagattgtgtgtgtgtgtgtgtgtgtgtgtgtgt 120
OY 121 GGTGACAGCATGGGAGGTTCACAGCTGATATGGGACCAAAATGCTAAGTGGG 180
DB 121 ggtgacaagcagcgagagttcaacactgtacatgacgacacaaatgctgaagtgcg 180
OY 181 ACGATTTATTCAGTTAGAGCCATGTACTGTTTGGTTCAACACCACTGTTTAAAC 240
DB 181 acgattattatcgttagagcgcattgtcatttgttcaacaccaccctgttacc 240
OY 241 AAAATTATGATTTAGATACCAAAAGTATGATACGCAAAAGACATTCGTTAGAC 300
DB 241 aaaattattgttgaagaatgcaaaagtagatacgcaaaaagacattcgttagac 300
OY 301 ACACATGGTATGATGAGACGTTACGGGCTCAGGTATGAGTTGATGCGTTGATGAA 360
DB 301 aacatggtgatagagacgttaccggtcaggtatgagttgaaatgctgtatgaa 360
OY 361 CGGCGATCACAATTTACCAAGACTTTTAAAGCAAAAGCAACCACTGCCACAAAT 420
DB 361 cggcgatcacaaattatcacagactttttaaagcaaaagcaaacactgcccacaaatt 420
OY 421 ACAGTGAAGTATCTGCAAGTTAGATGCTAAACAGCAACCAATGATTAAGACAAAGTCAA 480
DB 421 acagtgaagtatctgcgaagtttagatgtlaaacaagcaatgataatgacaagaatcaa 480
OY 481 TGGATTACTAAAGAGGTTAAACAGATGCTAATTAAGTAAAGATGACAGACAGCA 540
DB 481 tggattactaaagagtttaacaagatgctataaagtaagacatcgacaagcaga 540
OY 541 GTGTAACTGAGAGAGTACAGTTGAATAGATGCCAATATTAATACAGTATTCAA 600
DB 541 gtgttaactgagagagcgtacagttgaatagatgccacaataacacagatttcaa 600
OY 601 GATGAAAAAACCCATTAATAAATATTTGCTAAAGTCTGGAAATATTCATTATTAACAG 660
DB 601 gatgaaaaaacccctataaagtaatatgtctaagctcgggaatatatttcaatcag 660

OY 661 CAAATTTATCAAGATGATCAACACCAATTTGATATATACGAAATCCAAATTTAACA 720
DB 661 caaatttatacaagatgataccaacccaatttgatatactatctgaaatccaatttaca 720
OY 721 AGCAATCAACACATATTGAATATTATTTACTGAAAGTCTTGTGATTTAACACAAATCTT 780
DB 721 agcaatcaaacacatattgaattattacttgagacttgagcttggttlaacaacaaattctt 780
OY 781 CACAATTTATTAAGAGAGGTTGGAAGTCTTGCTACTGACGCGAGTCCAAACCACTACT 840
DB 781 cacaatttataaagaagaggttggaacttgcagctgagcgaggtccaacacactact 840
OY 841 TCAGAAATTCCTCATCTATTATATAGATGAATTTTCTCTATTATTCGCCGAAATTAAT 900
DB 841 tcagaatttcctcatctattatataagatgaattattctcattatgcccgaattatc 900
OY 901 GCGGATCTGGAATTTATCAATTTTATCAACCAAAATGATGATGATGATGATGCG 960
DB 901 gcggatctggaatttatcaatttattcaaaacaatgatgtgattgagataccagatcg 960
OY 961 AACCAATTTGAATTTGTTCAATTCGAGTTTAAATCAAAATGTTAATTAACCTTACGA 1020
DB 961 aaccaatttgaatttgcattccgagttatcaaaatgtaaatcaacttaca 1020
OY 1021 AAGAAGTGA 1029
DB 1021 aagaagtga 1029

RESULT 2
X59913
ID X59913 standard; DNA; 3336 BP.
XX
AC X59913;
DT 04-AUG-1999 (first entry)
XX
DE Riboflavin biosynthesis operon containing the ribg gene.
XX
KW Pyrimidine deaminase; pyrimidine reductase; ribg; antagonist;
KW agonist; antimicrobial; antibiotic; Helicobacter pylori infection;
KW H. pylori-induced cancer; gastrointestinal carcinoma; gastric ulcer;
KW gastritis; ss.
XX
OS Staphylococcus aureus.
XX
XX W09926475-A1.
PN 03-JUN-1999.
PD
XX
PE 23-NOV-1998; 98WO-US24857.
XX
PR 25-NOV-1997; 97US-0978456.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI Black MT, Burnham MKR, Fedon JC, Hodgson JE, Knowles DJC;
PI Lometto MA, Nicholas RO, Palmer LM, Pratt JM, Reichard RW;
PI Rosenberg M, Tiran CM, Ward JM, Warren RL;
XX
DR WPI: 1999-347572/29.
XX
PT New Staphylococcus aureus polypeptide and polynucleotide useful in
XX the treatment of gastric ulcer and gastritis
PS
XX Example 2; Page 30-31; 48pp; English.
XX
CC The present sequence represents the Staphylococcus aureus
CC riboflavin biosynthesis operon containing the pyrimidine
CC deaminase and pyrimidine reductase (ribg) gene. Staphylococcus
CC aureus ribg and its antagonists are used to treat individuals in
CC need of them. Disease related to expression or activity of ribg can

CC be determined by analysing the nucleic acid sequence encoding ribG
 CC or detecting the ribG polypeptide in a sample. ribG can also be used
 CC to identify antagonists or agonists. ribG, or its related nucleic acid,
 CC also has use as a vaccine to induce an immunological response in an
 CC animal. Antimicrobial compounds (e.g. agonists and antagonists of ribG),
 CC especially broad-spectrum antibiotics, may be of use in the treatment
 CC of *Helicobacter pylori* infection. This should decrease the advent of
 CC H. *pylori*-induced cancers, such as gastrointestinal carcinoma. The
 CC treatment should also cure gastric ulcers and gastritis.

XX Sequence 3336 BP; 1202 A; 460 C; 677 G; 997 T; 0 other;

Query Match 100.0%; Score 1029; DB 20; Length 3336;
 Best Local Similarity 100.0%; Pred. No. 4, 9e-230;
 Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 ATGATTATCGATTCACTTGCATAATATGTACAAAGTCAACAGGTGTTATTCACCC 60
    |||||||
DB 1 atgattatcgattcacttgcataatgtacaaaggtgttatccacc 60
OY 61 GTTGCGCGCTGTAGTTAATGAAGTAGAATGTTGTTGTCACACTTGACAAA 120
    |||||||
DB 61 gtcgcgctgtagttaatgaagtagaattgttattgtgcacacttgaagaaa 120
OY 121 GGTGACAGCATGCGGAGTTCAAGCACTGTATGCGACAAACAATGCTGAAGTGGC 180
    |||||||
DB 121 ggtgacagcatgcgagttcaagcactgtatgacaaacaatgctgaagtcgc 180
OY 181 ACATTTAATTAATAGTAGAGCATGATGATTTGTTGTTCAACACACCTGTATAC 240
    |||||||
DB 181 acatttatctacgttaagccatgtagcatcttgttccaaaccctgtgttaac 240
OY 241 AAAATTATGTTGTAAGATAGCAAAAAGTATAGCAACAAAACAATTCGTTAGAC 300
    |||||||
DB 241 aaattatgtatgtaagatagcaaaaagtatagcaacaaaacaatctgttagac 300
OY 301 ACACATGTTGATGAGAGCTTACGCGGTACGGTATGAGTGTGATGATGATGAA 360
    |||||||
DB 301 acacatggtgatgagagcttaccggtacgggtatgagtgagtcgttgaagaa 360
OY 361 CGGCGATCACAATTTATACAGACTTTTAAAGCAAAACAAAGCAACCTGCAATTT 420
    |||||||
DB 361 cggcgatcacaaatTTATACAGACTTTTAAAGCAAAACAAAGCAACCTGCAATTT 420
OY 421 ACGTGAAATATCTGCAAGTTTATGATGTAACACGATGATTAATGCAACAAAGTCAA 480
    |||||||
DB 421 acgtgaaatctgcaagttttagatgtaacaaagcgaatgataatgacaagtc 480
OY 481 TGGATTACTACAAAGGTTAAACAAGATGCTATAGTAAAGACATCGACACGACGA 540
    |||||||
DB 481 tggattactacaaagggttaacaagatgctataagtaagacacgacagagca 540
OY 541 GTGTTAACTGGAAGACGTATGTAATGATGATCCACAATATATACATGATTTCAA 600
    |||||||
DB 541 gttgtaactggaagacgtatgtaattgatgataatccacaatactacagattcaa 600
OY 601 GATGGAAGAAAACCTATATAAGTAATTTGTCTGAGTCTGGGAATATTTCAATTCAG 660
    |||||||
DB 601 gatggaagaaaacctataaagaataatgtctaagtcgtggaatacttcaacag 660
OY 661 CAATTTATCAAGATGAACAACCAACCAATTTGATATCTGAAATTCGAATTTTACA 720
    |||||||
DB 661 caatttatcaagatgaacaacccaatttgatcatatactgaaatccaaatttaca 720
OY 721 AGCAATCAACACATATTTGAATTTATTTACTTGAAGTCTTGATTTAACAACATTTCT 780
    |||||||
DB 721 agcaatcaaacacatattgaattattacttgaagtcgtgtgttttaacaacattc 780
OY 781 CACAATTTATATAAGAGAGATTGGAAGTTTGTCTAGTCGAGCAGGTCCACCACTACT 840
    |||||||
DB 781 cacaatttatataaagaagattggaacttgtctagtcgagcaggtccaaacctact 840

```

```

OY 841 TCAGAAATTCATCTATATATAGATGAATTTATTTCTATTATGCCCGGAATTAAT 900
    |||||||
DB 841 tcgaatttcctcatcatatataatagatgaatttattctattatcccggaattaat 900
OY 901 GCGGATCTGGAATTTATCAATTTTCAACAAGATGTATTAGATACAGATGCG 960
    |||||||
DB 901 gcggatctggaatttatcaatTTTCAACAAGATGTATTAGATACAGATGCG 960
OY 961 AACCAATTTGAATTTGTTCAATTCGAGTTATTAATCAAAAATGTTAAATTAATTACGA 1020
    |||||||
DB 961 aaccaatttgaatttgcattccaggttataatcaaaaatgttaatttaacttaca 1020
OY 1021 AAGAGTGA 1029
    |||||||
DB 1021 aagaagtga 1029

```

```

RESULT 3
X77007
ID X77007 standard; DNA; 3336 BP.
XX
XX X77007;
XX
XX 06-AUG-1999 (first entry)
DT
XX
XX S. aureus riboflavin biosynthesis operon.
DE
XX GTP cyclohydrolase II: ribA; diagnosis: vaccine; bacterial adhesion;
KW Helicobacter pylori infection; gastrointestinal carcinoma; gastric ulcer;
KW gastritis; wound healing; infection; drug screening; therapy;
KW riboflavin biosynthesis operon; ss.
XX
XX Staphylococcus aureus.
OS
XX
XX WO926646-A1.
XX
XX 03-JUN-1999.
PD
XX
XX 24-NOV-1998: 98WO-US25068.
XX
XX 25-NOV-1997: 97US-0977554.
PR
XX
XX (SMK ) SMITHKLINE BEECHAM CORP.
XX
XX Black MT, Burnham MKR, Debouck C, Fedon JC, Hodgson JB;
PI Jaworski DO, Knowles DJC, Lonetto MA, Mooney J, Nicholas RO;
PI Palmer LM, Pratt JM, Reichard RW, Rosenberg M, Traini CM;
PI Wang M, Ward JM, Warren RL, Zhong Y;
XX
XX MPI: 1999-357736/30.
XX
XX New Staphylococcus aureus ribA polypeptide - and methods for using
PT the polypeptide to screen for antibacterials
PS
XX
XX Claim 24; Page 36-37; 53pp; English.

```

CC This sequence represents the *S. aureus* riboflavin biosynthesis operon,
 CC which contains the ribA coding sequence of the invention. ribA belongs to
 CC the GTP cyclohydrolase II family. The polynucleotides and polypeptides
 CC may be employed as research reagents and material for the discovery of
 CC treatments and diagnostics for diseases, particularly human diseases.
 CC They can be used for diagnosis of the disease and staging of disease, and
 CC as reagents in differential screening methods. The polynucleotides may be
 CC used as a source for hybridisation probes, and for screening of genetic
 CC mutations, serotype, and identification, and for organism chromosome
 CC identification. The polypeptides can be used to produce antibodies. The
 CC polypeptides can also be used in vaccine formulations, and to identify
 CC agonists and antagonists. These are used to prevent, inhibit or treat
 CC diseases, particularly of *Helicobacter pylori* infections, such as
 CC gastrointestinal carcinoma. Gastric ulcers and gastritis may also be
 CC treated. The polypeptides can also be used to treat wounds and
 CC in-dwelling devices to prevent bacterial adhesion and infection, and to
 CC block ribA protein-mediated mammalian cell invasion. The frequency of

Staphylococcal infections has risen dramatically, and it is no longer common to find isolated *S. aureus* strains that are resistant to standard antibiotics. The rDNA products of the invention can be used screen for new antibacterial compounds that may target these resistant bacteria.

Sequence 3336 BP; 1202 A; 459 C; 678 G; 997 T; 0 other;

Query Match 100.0%; Score 1029; DB 20; Length 3336;
Best Local Similarity 100.0%; Pred. No. 4.9e-230;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGATTTATGCGATTCACCTGCAATATGTACAGAGTCAACAGCTGTTATCCACC 60
DB 1 atgattatgcttcaacttgcataatgtgtaacaggtcaacagtggttattccacc 60
QY 61 GTGGCGCTGTTGTGTTAATAGTAGAGTTGTTGGATTGGTGCACACTTGAAGAAA 120
DB 61 gtggcgctgtgtgttaattagagtgagtggtgtgtgtgtgtgtgtgtgtgtgtgt 120
QY 121 GTTGACAGCAGCAGCGAGGTTCAACACTTGATATGACACAAATGCTGAGAGTGG 180
DB 121 gttgacagcagcagcgaggttcaacacttgatattgacacaaatgctgagagtggtg 180
QY 181 ACGATTATATTACGTTAGAGCCATGTAGTCATTTTGGTTCAACACACCCCTGTGTAAC 240
DB 181 acgattatattacgttagagccatgttagcttltgtgttcaacacccctgtgttaac 240
QY 241 AAAATTATGATTGTAAGTAGCAAAAGTAGTATACGCACAAAGACAAATTCGTTAGC 300
DB 241 aaattatgtatgtaagtagcaaaagtagtatacgacaaagacaaattcggttagc 300
QY 301 ACACATGCTGATGAGACGCTTACGCGCTCAGCGTATTGAGTGAAGTGGCTGATGAA 360
DB 301 acacatgctgatgagacgcttaccgctcagcgctatgagtgatgagtgatgagtaa 360
QY 361 CGGCATCACAATTATACCAAGACTTTTAAAGCAAAAGCAACCACTGCCACAATT 420
DB 361 cggcatcacattataccagactttttaaagcaaaagcaaacactgcccacaatt 420
QY 421 ACAGGAAGATCTGCAAGTTTATGTAACAAGCAAGATGATGAGAAAGTCAA 480
DB 421 acaggaagatctgcagagtttagtgaacaagcaagcaaatggaagaagtgcaa 480
QY 481 TGGATTACTAACAAGAGGTTTAAACAGATCTTAACTTAAAGATGACAGACGCA 540
DB 481 tggattactaacaagaggttaacaagatgcttaagttgaacatcgacaagcga 540
QY 541 GTGTTAAGTGAAGAGCTGACGTTGATGATCCACAATATACGATTCACGATTC 600
DB 541 gttttaagtggaagagctgacgttgattagatgccacatatatacagatttcaa 600
QY 601 GATGAATAAACCCCTATAAATATATTTGCTAGTCTGGAAATATTCATTTATACG 660
DB 601 gatgaataaacccctataaataatattgtctagtcgaggaatattcaattcaag 660
QY 661 CAAATTTTCAAGATGATCAACACCAATTTGATATATACGAATAATCAATTTAACA 720
DB 661 caaatTTTCAAGATGATCAACACCAATTTGATATATACGAATAATCAATTTAACA 720
QY 721 AGCAATCAACAATATTGAAATTTACTTGAAAGCTGTGATTTAACAACAATTCCT 780
DB 721 agcaatcaacaatattgaaatttacttgaaagctgtgatttaacaacaattcct 780
QY 781 CACATTTATATTAAGAGAGTGGAGCTTGTGATGAGAGGAGGTCACCACTACT 840
DB 781 cacatttatataaagagagtgagacttgcagtcgagagcaggtccacccact 840
QY 841 TCAGAAATTCCTATATATATATAGATTAATTTCTATATAGCCCGCAATTAAT 900
DB 841 tcagaaatTCCTATATATATATAGATTAATTTCTATATAGCCCGCAATTAAT 900
QY 901 GCGGATCTGGAATTTATCAATTTTATCAAAACAATGATGATTGAGATACCAATGCG 960

```

```

DB 901 ggcgacttggaattatcaatttatacaacaatgatgattgagatccagatcg 960
QY 961 AACCAATTTGAAATGTCATCCGAGTTATTAATCAAAATGTTAATTAATTTACGA 1020
DB 961 aaccattgaaattgtcatccaggttataatcaacaatgtaatttaacttaaga 1020
QY 1021 AAGAAGTGA 1029
DB 1021 aagaagtga 1029

RESULT 4
ID X60277 standard; DNA; 3336 BP.
XX
AC X60277;
XX
DT 11-AUG-1999 (first entry)
XX
DE Riboflavin biosynthesis operon derived open reading frame.
XX
KW Riboflavin synthase; alpha-subunit; ribB; vaccine;
KW Helicobacter pylori infection; gastrointestinal carcinoma;
KW gastric ulcer; gastritis; bacterial adhesion; infection;
KW cell invasion; bacteriostatic; bactericidal; ss.
XX
OS Staphylococcus aureus.
XX
FN W09926650-A1.
XX
PD 03-JUN-1999.
XX
PF 24-NOV-1998; 98WO-US25070.
XX
PR 25-NOV-1997; 97US-0977866.
XX
PA (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI Black MT, Burnham MKR, Fedon JC, Hodgson JE, Knowles DJC;
PI Loretto MA, Nicholas RO, Palmer LM, Pratt JM, Reichard RW;
PI Rosenberg M, Ttranl CM, Ward JM, Warren RL;
XX
DR WPI: 1999-370857/31.
XX
PT New Staphylococcus aureus ribB polypeptide, useful in the prevention
XX and treatment of Helicobacter pylori infections
XX
PS Example 2; Page 30-31; 44pp; English.
XX
CC The present sequence represents an open reading frame of the
CC riboflavin biosynthesis operon, containing the ribB sequence. RibB is a
CC riboflavin synthase (alpha-subunit) protein. It is isolated from
CC Staphylococcus aureus. The polynucleotides and polypeptides may be
CC employed as research reagents for diseases, particularly human diseases.
CC treatments and diagnostics for diseases, particularly human diseases.
CC They can be used for diagnosis of the disease and staging of disease,
CC and as reagents in differential screening methods. The polynucleotides
CC may be used as a source for hybridization probes, and for screening
CC of genetic mutations, serotype, and identification, and for organism
CC chromosome identification. The polypeptides can be used to produce
CC antibodies. The polypeptides can also be used in vaccine formulations,
CC and to identify agonists and antagonists. These are used to prevent,
CC inhibit or treat diseases, particularly of Helicobacter pylori
CC infections, such as gastrointestinal carcinoma. Gastric ulcers and
CC gastritis may also be treated. The polypeptides can also be used to
CC treat wounds and in-dwelling devices to prevent bacterial adhesion and
CC infection, and to block ribB protein-mediated mammalian cell invasion.
CC The ribB polypeptide can also be used to screen for antagonists and
CC agonists which may be bacteriostatic and/or bactericidal, and so may
CC also be employed to inhibit and treat diseases.
XX
SO Sequence 3336 BP; 1202 A; 459 C; 678 G; 997 T; 0 other;

```

Query Match 100.0%; Score 1029; DB 20; Length 3336;
 Best Local Similarity 100.0%; Pred. No. 4.9e-230;
 Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 ATGATTTATCGATTCACTTGCATAATATGTTCAAGTCAACAGGTGTTAAATCCACC 60
    |||
DB 1 atgattatcgattcacttgcataatggtacaaggtcaacaggtgttaatccacc 60
OY 61 GTTGGCGCTGTTAGTTAAATGAAGTAGATTGTTGTTGTTGTCACACTTGAGAAA 120
    |||
DB 61 gttgcgcgtgttagttaaatgaagtagattgttggttgttcacacttgagaaa 120
OY 121 GTTGACAAGCATGCGGAGTTCAAGCCTTGATATGTCACAACAAATGCTGAAGTGGC 180
    |||
DB 121 gttgacaagcatgcgagttcaagccttgatattgacacacaacaaatgctgaagtggc 180
OY 181 ACGATTATATACGTAGAGCCATGATGATTTGTTCAACACCCCTGTGTAC 240
    |||
DB 181 acgattatatcgttagagccatgtagcatttggttcaacaccctgtgtaac 240
OY 241 AAAATTATGTTAGTAAAGATGCAAAAAGTAGTACGCAACAAAGCAATTGCTTAGAC 300
    |||
DB 241 aaattattgattgtaagatagcaaaaagtagtatacgcaaaaagcaattcgttagac 300
OY 301 ACACATGGTGATGAGACGTACGGGCTCAGGTTATTGAGTTGATCGCTTGATGAGAA 360
    |||
DB 301 acacatggtgatgagacgtacgggctcagggtattgagttgattcgcttgatgagaa 360
OY 361 CGGCGATCAACAATTATACAGACTTTTAAAGCAAAACAAAGCAACGCAACAATT 420
    |||
DB 361 cggcgatcaacaattatacagactttttaaagcaaaaacaaagcaacgcgcacaatt 420
OY 421 ACAGTGAAGTATCTGCAAGTTTATGATGTTAAACAGCGAATGATTAATGACAAAGTCAA 480
    |||
DB 421 acagtgaagtatctgcaagttttagatgtttaaacagcgaaatgatttaattgacaa 480
OY 481 TGATATTACTACAAGAGGTTAAACAGATGCTATAGTTAGACATCGACACGACGCA 540
    |||
DB 481 tgattactacaagagggttaaacaagatgctataagttagacatcgacacgacgca 540
OY 541 GTTAACTGAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
    |||
DB 541 gtttaactgagacgattgattgattgattgattgattgattgattgattgattgattgatt 600
OY 601 GATGAAAAAACCCCTATATAAGTATGTTGCTAAGTCTGGGAATATTCATTTTAAATCAG 660
    |||
DB 601 gatgaaaaaacccctataaagtaattgtctaagtcctgggaattattcattttaaaccag 660
OY 661 CAATTTTATCAAGTGAATACACCAATTTGATATATCTGAAAATCCAAATTTTACA 720
    |||
DB 661 caatttatcaagtgaatacaacccaatttgatataatctgaaaatccaatttaca 720
OY 721 AGCAATCAACACATATTTGAATTTATTTACTTGAAGCTTGATTAACAAACAATCTT 780
    |||
DB 721 agcaatcaaacacatattgaattatttacttgaagcttgatttaacaacaattctt 780
OY 781 CACAATTTATATAAAGAGAGTGTGAACCTTGCTAGTCAGGACGAGTCCACCACTACT 840
    |||
DB 781 cacaatttatataaagagagtgtagacttgcagaggttgcagaggttccaaccactact 840
OY 841 TCGAATTCCTCATCTTATATAGATGAATTTATCTCTATATATGCGCCGGAATTAAT 900
    |||
DB 841 tcgaatttcctcatcttattatagatgaatttatctctatattatgccccgaatttaatt 900
OY 901 GGGGATCTGGAATTTATCAATTTTATCAACAATGATGATGATGATGATGATGATGATGAT 960
    |||
DB 901 ggggatctggaatttatcaattttatcaacaatgatgatgatgatgatgatgatgatgatgat 960
OY 961 AACCAATTTGAAATTTGTTCAATTCGAGTTTAAATCAAAATGTTAAATTAATTTACGTA 1020
    |||
DB 961 aaccaatttgaatttgtctatccgagtttaatacaaaatgtttaatttaacttaccga 1020
  
```

OY 1021 AAGAGTGA 1029
 DB 1021 aagaagtga 1029

RESULT 5
 X60830
 ID X60830 standard; DNA; 3336 BP.
 XX
 AC X60830;
 XX
 DT 09-AUG-1999 (first entry)
 XX
 DE S. aureus riboflavin biosynthesis operon.
 XX
 KW Lumazine synthase; ribH; riboflavin; differential screening; gastritis;
 KW H. pylori infection; gastrointestinal carcinoma; gastric ulcer; wound;
 KW bacterial adhesion; cell invasion; riboflavin biosynthesis operon; ss.
 XX
 OS Staphylococcus aureus.
 XX
 PN W0927128-A1.
 XX
 PD 03-JUN-1999.
 XX
 PF 24-NOV-1998; 98MO-US25096.
 XX
 PR 25-NOV-1997; 97US-0977553.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PI Debowick C, Fedon JC, Jaworski D, Mooney J, Palmer LM;
 PI Tsalni CM, Wang M, Warren RL, Zhong Y;
 DR WPI: 1999-357850/30.
 XX
 PT New Staphylococcus aureus ribH polynucleotide and polypeptide useful
 PT in the treatment of gastric ulcers and gastritis
 XX
 PS Example 2; Page 34-36; 48pp; English.
 XX

The invention relates to polynucleotides and polypeptides of the lumazine synthase (beta -subunit) family, designated ribH. They are isolated from Staphylococcus aureus. ribH enzymes are involved in the synthesis of riboflavin. The polynucleotides and polypeptides may be employed as research reagents and material for the discovery of treatments and CC diagnostics for diseases, particularly human diseases. They can be used CC for diagnosis of the disease and staging of disease, and as reagents in CC differential screening methods. The polynucleotides may be used as a CC source for hybridization probes, and for screening of genetic mutations, CC serotype, and identification, and for chromosome identification. The CC polypeptides can be used to produce antibodies. The polypeptides can CC also be used in vaccine formulations, and to identify agonists and CC antagonists. These are used to prevent, inhibit or treat diseases, CC particularly of Helicobacter pylori infections, such as gastrointestinal carcinoma. Gastric ulcers and gastritis may also be treated. The CC polypeptides can also be used to treat wounds and in-dwelling devices to CC prevent bacterial adhesion and infection, and to block ribH protein- CC mediated mammalian cell invasion. Antagonists and agonists of ribH may CC also be employed to inhibit and treat diseases. The present sequence CC represents the S. aureus riboflavin biosynthesis operon that encodes CC genes ribG, ribB, ribA and ribH.
 CC
 XX
 SQ Sequence 3336 BP; 1202 A; 459 C; 678 G; 997 T; 0 other;

Query Match 100.0%; Score 1029; DB 20; Length 3336;
 Best Local Similarity 100.0%; Pred. No. 4.9e-230;
 Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 ATGATTTATCGATTCACTTGCATAATATGTTCAAGTCAACAGGTGTTAAATCCACC 60
    |||
  
```

Db 1 atgatatatgcattcaacttgcataatagtacaaagtcacaaacagtgcttaaccacc 60
Oy 61 GTTGCGCGTGTGTGTAATGAAGTAGATTGTGATTTGGTGCACACTTGSAGAAA 120
Db 61 gtggcgctgtgttgatcaatgaagtgatctgtctatgtgtcgaacttgaagaaa 120
Oy 121 GGTGACAGCATGCGAGGTTCAACACTTGTATATGCGACAACAATGCTGAAGTGGC 180
Db 121 ggtgcaagcatgcgaggttcaagcacttgatgctgcacacaatagtctgaagtgcg 180
Oy 181 ACGATTTATATTACCTTAGAGCCATGTACTCAATTTGGTTCAACACCACCTGTGTAA 240
Db 181 acgattatattacgttagagccatgtactcttgtgttcacacccaccgtgttaac 240
Oy 241 AAAATTATGATTGTAAAGTACCAAAAGTAGATATCGCAACAAAAGACAATGTTAGAC 300
Db 241 aaaattatgtatcgaatgaacaaagatgatacgcacaaagacatcgtttagac 300
Oy 301 ACACATGCTGATGAGACGTTTACGGGCTCAGCGTATTTGAAGGTTGAATGCGTTGATGAA 360
Db 301 acacatgctgtatgagacgctttagcggtcagcgctatgtggttgaatgctgtatgaa 360
Oy 361 CGGCATCACAAATTATACCAAGACTTTTAAAGCAAAAAGCAAGCACTGCCACAATT 420
Db 361 cgggcatacaattataccaagacttlttaagcaaaagcaaacactgcacaact 420
Oy 421 ACAGTGAAGTATCTGCAAGTTAGTGTAAACAAGCAATGATATAGTACAAAGTCAA 480
Db 421 acagtgaagatctcgcagctttagatgtaacaagcaagcaatgaatgaagcaaac 480
Oy 481 TGGATTACTTAACAAAGAGTTTAAACAGATGCTATTAAGTTAAGACATGACACGACGA 540
Db 481 tggattacttaacaaagaggttaacagatgtctaaagtaagatcagacacagca 540
Oy 541 GTGTTAACTGGAAGCGTACATGATTAAGTATCCATATACTACACTATTCATA 600
Db 541 gtgttaactggaagcgttaacagttgaatagatgtaccacaatactacacgtattcaa 600
Oy 601 GATGAAAAAACCTATATAAAGTATATGCTAAGTGGAGATATTCATTTAATACG 660
Db 601 gatgaaaaaacctataaagtaataatgtcttaagtcggaatattcaattcaacg 660
Oy 661 CAATTTATCAAGATGAATCAACACCAATTTGATATATATCAAAATTCAAATTAACA 720
Db 661 caattatcaagatgaatcaacacacatlttgatataaacgaaatccaaatttaaca 720
Oy 721 AGCAATCAACACATATTGAATTTATTACTTGAAGCTCTTGATTTTAAACAATTCCT 780
Db 721 agcaatcaaacacatatgaattatlttacttgaagcttctgtatlttaacacaaattc 780
Oy 781 CACATTTATATAAAGAGGAGTTGAACTTGCTAGTGAAGCGAGTCCAACTACT 840
Db 781 cacaattatataaagagagtggaacttgcctagtcgagcggtccaaactact 840
Oy 841 TCAGATTCTCCATCTATTATATAGATGAATTTCTCTATTAATGCCCGAAATTAAT 900
Db 841 tcagattctccatctatataatagatgaattatctctatataagcccgaaatla 900
Oy 901 GCGGATCTGGAATTAATCAATTTTATCAACAAATGATGATTTGAGATACCAGATGCG 960
Db 901 gcggatctggaaatataatcaattatcaacaatgattgagttacagatgag 960
Oy 961 AACCAATTTGAATTTGTTTCATTCGAGTATATAATCAAAATGTTAAATTAATCTTACA 1020
Db 961 aaccatttgaatattgtatcctcgagttatcaaatcaaatglttaacttaacttaca 1020
Oy 1021 AAGAAGTGA 1029
Db 1021 aagaagtga 1029

ID V74453 standard; DNA; 7588 BP.
XX
AC V74453;
XX
DT 16-MAR-1999 (first entry)
XX
XX
DE Staphylococcus aureus config SEQ ID #142.
XX
KW Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW skin infection; surgical wound infection; scalded skin syndrome;
KW toxic shock syndrome; ds.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT misc-feature 651..720
FT /*tag= a
FT /note= "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"
FT misc-feature 2461..2520
FT /*tag= b
FT /note= "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"
FT misc-feature 4261..4320
FT /*tag= c
FT /note= "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"
FT misc-feature 6061..6120
FT /*tag= d
FT /note= "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"
PN EP786519-A2.
XX
XX 30-JUL-1997.
XX
PE 07-JAN-1997; 97EP-0100117.
XX
PR 05-JAN-1996; 96US-0009861.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Barash SC, Chol GH, Dillon PJ, Fannon MR, Kunsch CA;
PI Rosen CA;
XX
XX WPI: 1997-374922/35.
XX
DR Polynucleotide(s) and proteins derived from Staphylococcus aureus
XX stored on computer readable medium and used in the production of
XX anti-S.aureus vaccines
XX
XX
XX Claim 1; Page 740-744; 3271pp; English.
XX
CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences
CC of the invention. The DNA sequences are recorded on a computer readable
CC medium, preferably selected from a floppy or hard disk, random access
CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC the S.aureus DNA sequences allows putative functions to be assigned so
CC that protein-encoding or regulatory regions of commercial, therapeutic or
CC industrial importance can be obtained. Specifically, sequences which are
CC likely to encode antigens have been identified and these polypeptides can
CC be used in a vaccine composition against S.aureus infection. The
CC polypeptides can also be used in a kit for the immunodetection of
CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,

including cellulitis, eyelid infections, food poisoning, osteomyelitis, skin and surgical wound infections, scaled skin syndrome, toxic shock syndrome, etc. Organisms transformed with the DNA sequences can be used for recombinant production of the polypeptides. The new DNA sequences (and their fragments) are useful as primers or probes for isolating homologues of any of the *S. aureus* DNA sequences contained on the computer readable medium.

Sequence 7588 BP; 2333 A; 1341 C; 1061 G; 2613 T; 240 other:

Query Match 97.2%; Score 1000; DB 18; Length 7588;
Best Local Similarity 98.7%; Pred. No. 3,3e-223;

Matches 1019; Conservative 0; Mismatches 10; Indels 3; Gaps 1;

```

Oy 1 ATGGATTATGCGATTCACCTTGCACAAATATGGTACAAAGTCAAAAGGTGTTAATCCACC 60
    |||||||
Db 5849 ATGGATTATGCGATTCACCTTGCACAAATATGGTACAAAGTCAAAAGGTGTTAATCCACC 5790
Oy 61 GTTGGGCGCTGTAGTATGAAGTAGATTTGTTGTTGTCACACTTGAGAAA 120
    |||||||
Db 5789 GTTGGGCGCTGTAGTATGAAGTAGATTTGTTGTTGTCACACTTGAGAAA 5730
Oy 121 GTGACACAGCATCGAGGTTCAAGCACTGATATGGACACACAAATGCTGAAGTGG 180
    |||||||
Db 5729 GTGACACAGCATCGAGGTTCAAGCACTGATATGGACACACAAATGCTGAAGTGG 5670
Oy 181 ACGATTATATTCAGTTAGAGCCATGATGATTTGGTTCAACACCCCTGTGTTAC 240
    |||||||
Db 5669 ACGATTATATTCAGTTAGAGCCATGATGATTTGGTTCAACACCCCTGTGTTAC 5610
Oy 241 AAAATTATGATTGTAAAGATAGCAAAAGTAGTATAGCAACAAAGCATTTGTTAGAC 300
    |||||||
Db 5609 AAAATTATGATTGTAAAGATAGCAAAAGTAGTATAGCAACAAAGCATTTGTTAGAC 5550
Oy 301 ACACATGATGATGAGACGTTACGGCTCAGCATTTGAGGTTGAATGCGTTGATGAA 360
    |||||||
Db 5549 ACACATGATGATGAGACGTTACGGCTCAGCATTTGAGGTTGAATGCGTTGATGAA 5490
Oy 361 CGGCGATCACAATTAACCAAGACTTTTAAAGCAAAAGCAAAAGCAATGCGACAAT 420
    |||||||
Db 5489 CGGCGATCACAATTAACCAAGACTTTTAAAGCAAAAGCAAAAGCAATGCGACAAT 5430
Oy 421 ACGTGAAGTATCTGCAATTTAGATGTTAAACCAAGCATATATGACAAATGCA 480
    |||||||
Db 5429 ACGTGAAGTATCTGCAATTTAGATGTTAAACCAAGCATATATGACAAATGCA 5370
Oy 481 TGGATTACTAACAAAGGTTAAACAAGATGCTATAGTTAAGACATCGACGACGCA 540
    |||||||
Db 5369 TGGATTACTAACAAAGGTTAAACAAGATGCTATAGTTAAGACATCGACGACGCA 5310
Oy 541 GTGTTACTGGAAGACGTACAGTTGAATTAGATGATCCACAATATACATGATTTCAA 600
    |||||||
Db 5309 GTGTTACTGGAAGACGTACAGTTGAATTAGATGATCCACAATATACATGATTTCAA 5250
Oy 601 GATGAAAAAACCCCTATAAAGTAATTTGTCTAAGTCTGGGAATATCTATTTAATCAG 660
    |||||||
Db 5249 GATGAAAAAACCCCTATAAAGTAATTTGTCTAAGTCTGGGAATATCTATTTAATCAG 5190
Oy 661 CAATTTATCAAGATGAATACCAATTTGATATATCTGAAAAATCCAAATTTTACA 720
    |||||||
Db 5189 CAATTTATCAAGATGAATACCAATTTGATATATCTGAAAAATCCAAATTTTACA 5130
Oy 721 AGCAATCAACACATATTTGAATATTTACTGAAGTCTTGATATTTAACACAATTC 780
    |||||||
Db 5129 AGCAATCAACACATATTTGAATATTTACTGAAGTCTTGATATTTAACACAATTC 5070
Oy 781 CACAATTTATATAAAGAGAGTGGAACTTGTCTAGTCAGGACAGGTCCACCACTACT 840
    |||||||
Db 5069 CACAATTTATATAAAGAGAGTGGAACTTGTCTAGTCAGGACAGGTCCACCACTACT 5010
Oy 841 TCGAATTCCTCCATCT--TTATATAGATGAATTTATCTCTATTTATGCGCGAAATTA 897
    |||||||

```

```

Db 5009 TCGAATTTCTCCAAATCTAATTATATAGATGAATTTATCTCTAATATGCCCGAAATTA 4950
Oy 898 ATGGGCGGATCTCGAAATATATCAATTTTATCAACAAAGATGATTTAGATACAGAT 957
    |||||||
Db 4949 ATGGGCGGATCTCGAAATATATCAATTTTATCAACAAAGATGATTTAGATACAGAT 4890
Oy 958 GCGAACCAATTTGAATTTGTCATTCGAGTTATTAAATCAAAATGTTAAATTAACCTTA 1017
    |||||||
Db 4889 GCGAACCAATTTGAATTTGTCATTCGAGTTATTAAATCAAAATGTTAAATTAACCTTA 4830
Oy 1018 CGAAGAGAGTGA 1029
    |||||||
Db 4829 CGAAGAGAGTGA 4818

RESULT 7
ID T84023 standard; DNA; 1269 BP.
XX
AC T84023:
XX
DT 01-SEP-1998 (first entry)
XX
DE DNA encoding a Riboflavin biosynthesis protein RIBG.
XX
KW Staphylococcus aureus protein; ribozyme; antisense sequence; control;
KW Staphylococcal gene; regulatory element; bacterial gene expression;
KW vaccine; Staphylococcal infection; food poisoning; scaled skin syndrome;
KW toxic shock syndrome; ss.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT CDS 734..1033
FT /tag= a
PN MO9730070-A1.
XX
PD 21-AUG-1997.
XX
PE 19-FEB-1997; 97WO-US02318.
XX
PR 20-FEB-1996; 96US-0011888.
XX
PA (SMK ) SMITHKLINE BEECHAM CORP.
XX
PI Black MT, Burnham MK, Hodgson JE, Knowles DUC, Nicholas RO;
PI Pratt JM, Reichard RW, Rosenberg M, Ward JM;
XX
DR WPI: 1997-424969/39.
DR F-PSDB: W28084.
XX
PT Novel polypeptide(s) from Staphylococcus aureus strain WCUH29 - used
PT to isolate antimicrobial compounds, and in vaccines against S.
PT aureus infection
XX
PS Claim 9; Page 825; 989pp; English.
XX
CC The present sequence encodes a Staphylococcus aureus protein, that,
CC based on homology with an Bacillus subtilis protein, is believed to be a
CC Riboflavin biosynthesis protein RIBG. The present sequence was obtained
CC from a library of clones of S. aureus WCUH 29 in Escherichia coli. The
CC DNA sequence can be used in the construction of ribozymes and antisense
CC sequences to control the expression of Staphylococcal genes. The DNA
CC sequence is also useful as a source of regulatory elements for the
CC control of bacterial gene expression. The encoded protein may be used
CC to produce vaccines to enable a host to produce specific antibodies
CC with antibacterial action. These vaccines and antibodies would protect
CC a host against invasion by S. aureus, and conditions relating to
CC Staphylococcal infection, e.g. Staphylococcal food poisoning, scaled
CC skin syndrome, and toxic shock syndrome.
XX
SQ Sequence 1269 BP; 435 A; 188 C; 256 G; 371 T; 19 other:

```

Query Match 34.2%; Score 352.2; DB 18; Length 1269;
Best Local Similarity 88.4%; Pred. No. 5.1e-73;
Matches 473; Conservative 0; Mismatches 44; Indels 18; Gaps 8;

QY 1 ATGATTTATGCGATTCACTTGCATAATATGTACAAAGGTCAAAACAGGTGTTATCCACC 60
DB 734 atgattatgcatcgaatcccaattcgaataatgtacaaaggtcnaaagcgtgtatccacc 793

OY 61 GTTGCGCTGTTGTAGTAAATGACAGTATGTTGGTATGTTGTCACACTTGAGAAA 120
DB 794 gttagcgtctgttagtgaatgaagtagatgtgttattgtgtgacacacttgaagaaa 853

OY 121 GGTGCAAGCANTGGAGTTCAACACTTGATATGCAACAATAATGTGAAGTGC 180
DB 854 ggtgcaagcagtcggaagttcaagcactgtatgcaacaanaatgctgaagtcg 913

OY 181 ACGATTTATATACGTTGACGCAATGTATGTTGGTTCACACACCCTGTTTAC 240
DB 914 acgatttatatagcttagagccatgtatgttctgttcaaccaccctgtgttaac 973

OY 241 AAAATTATGATTGTAGATAGCAAAAGTAGTA-TACCCAAACAAAGCAATT-CGTTAG 298
DB 974 aaattatgtattgtaagatagcaaaagtagtattacnaaanaagacaattccgttag 1033

OY 299 ACACACAT-GGTGATGAGACGTTACGGGCTCACGGTATTGAGTTGATCGCTT----- 351
DB 1034 acacacatggtgatgagacgcttacggtctccacgctatttgaaagctgaattgcgttg 1093

OY 352 GATGATGACGGCGCATCAATTTATACCAAGACTTTT---AAAGCAAAAGCAAGCAA 408
DB 1094 gatgatgacggcgcatcaatltaccaaagactttttaaagcaaaagcaagcaaa 1153

OY 409 CTGCCAC--AAATTACACTGAAGATATC-TGCACTTTAGATGTAAACAAGCG--AAT 462
DB 1154 ctgcacaaataattacagtgaagttcttgaaagtttagatggtttaaacaagcgaa 1213

OY 463 GATATGACAAAGTCATGATTTACTTAACAAGAGTTAAACAAGATGTCTATA 517
DB 1214 gataatgacaagaatgcaatgattactaacaaagagttaaacaagatgctcata 1268

RESULT 8
X59910 standard; DNA: 1269 BP.
XX X59910;
XX 04-AUG-1999 (first entry)
XX
XX Pyrimidine deaminase and pyrimidine reductase (ribG) ORF sequence.
XX
XX Pyrimidine deaminase; pyrimidine reductase; ribG; antagonist;
XX agonist; antimicrobial; antibiotic; Helicobacter pylori infection;
XX H. pylori-induced cancer; gastrointestinal carcinoma; gastric ulcer;
XX gastritis; ss.
XX
XX Staphylococcus aureus.
XX
XX
XX MO9926475-A1.
XX
XX 03-JUN-1999.
XX
XX 23-NOV-1998; 98WO-US24857.
XX
XX 25-NOV-1997; 97US-0978456.
XX
XX (SMIK) SMITHKLINE BEECHAM CORP.
XX
XX Black MT, Burnham MKR, Fedon JC, Hodgson JE, Knowles DUC;
XX Lometto MA, Nicholas RO, Palmer LM, Pratt JM, Reichard RW;
XX Rosenberg M, Trani CM, Ward JM, Warren RL;

XX
DR WPI: 1999-347572/29.
DR P-PSDB: Y15914.
XX
PT New Staphylococcus aureus polypeptide and polynucleotide useful in
PT the treatment of gastric ulcer and gastritis
XX
XX Claim 24; Page 5-6; 48bp; English.
XX
XX The present sequence encodes a Staphylococcus aureus pyrimidine
XX deaminase and pyrimidine reductase (ribG) polypeptide. Staphylococcus
XX aureus ribG and its antagonists are used to treat individuals in
XX need of them. Disease related to expression or activity of ribG can
XX be determined by analysing the nucleic acid sequence encoding ribG
XX or detecting the ribG polypeptide in a sample. ribG can also be used
XX to identify antagonists or agonists. ribG, or its related nucleic acid,
XX also has use as a vaccine to induce an immunological response in an
XX animal. Antimicrobial compounds (e.g. agonists and antagonists of ribG),
XX especially broad-spectrum antibiotics, may be of use in the treatment
XX of Helicobacter pylori infection. This should decrease the advent of
XX H. pylori-induced cancers, such as gastrointestinal carcinoma. The
XX treatment should also cure gastric ulcers and gastritis.
XX
XX Sequence 1269 BP; 435 A; 188 C; 256 G; 371 T; 19 other;

Query Match 34.2%; Score 352.2; DB 20; Length 1269;
Best Local Similarity 88.4%; Pred. No. 5.1e-73;
Matches 473; Conservative 0; Mismatches 44; Indels 18; Gaps 8;

QY 1 ATGATTTATGCGATTCACTTGCATAATATGTACAAAGGTCAAAACAGGTGTTATCCACC 60
DB 734 atgattatgcatcgaatcccaattcgaataatgtacaaaggtcnaaagcgtgtatccacc 793

OY 61 GTTGCGCTGTTGTAGTAAATGACAGTATGTTGGTATGTTGTCACACTTGAGAAA 120
DB 794 gttagcgtctgttagtgaatgaagtagatgtgttattgtgtgacacacttgaagaaa 853

OY 121 GGTGCAAGCANTGGAGTTCAACACTTGATATGCAACAATAATGTGAAGTGC 180
DB 854 ggtgcaagcagtcggaagttcaagcactgtatgcaacaanaatgctgaagtcg 913

OY 181 ACGATTTATATACGTTGACGCAATGTATGTTGGTTCACACACCCTGTTTAC 240
DB 914 acgatttatatagcttagagccatgtatgttctgttcaaccaccctgtgttaac 973

OY 241 AAAATTATGATTGTAGATAGCAAAAGTAGTA-TACCCAAACAAAGCAATT-CGTTAG 298
DB 974 aaattatgtattgtaagatagcaaaagtagtattacnaaanaagacaattccgttag 1033

OY 299 ACACACAT-GGTGATGAGACGTTACGGGCTCACGGTATTGAGTTGATCGCTT----- 351
DB 1034 acacacatggtgatgagacgcttacggtctccacgctatttgaaagctgaattgcgttg 1093

OY 352 GATGATGACGGCGCATCAATTTATACCAAGACTTTT---AAAGCAAAAGCAAGCAA 408
DB 1094 gatgatgacggcgcatcaatltaccaaagactttttaaagcaaaagcaagcaaa 1153

OY 409 CTGCCAC--AAATTACACTGAAGATATC-TGCACTTTAGATGTAAACAAGCG--AAT 462
DB 1154 ctgcacaaataattacagtgaagttcttgaaagtttagatggtttaaacaagcgaa 1213

OY 463 GATATGACAAAGTCATGATTTACTTAACAAGAGTTAAACAAGATGTCTATA 517
DB 1214 gataatgacaagaatgcaatgattactaacaaagagttaaacaagatgctcata 1268

RESULT 9
Q10165
ID Q10165 standard; DNA: 5567 BP.
XX
XX Q10165;
XX


```

XX DE B. subtilis rib operon nucleotide sequence.
XX KM Riboflavin; open reading frame; ORF; structural gene; promoter;
XX KM vitamin B2; Bacillus subtilis; rib operon; ds.
XX OS Bacillus subtilis.
XX FH Key
XX FH CDS
XX FT 1..5567
XX FT /tag= a
XX FT /note= "reading frame 1 translated protein (Y21801)"
XX FT 2..5567
XX FT /tag= b
XX FT /note= "reading frame 2 translated protein (Y21802)"
XX FT 3..5567
XX FT /tag= c
XX FT /note= "reading frame 3 translated protein (Y21803)"
XX PN US592538-A.
XX PD 20-JUL-1999.
XX PF 24-AUG-1998; 98US-0138775.
XX PR 11-SEP-1990; 90US-0581048.
XX PR 22-JUN-1989; 89US-0370378.
XX PR 21-APR-1992; 92US-0873572.
XX PR 06-FEB-1995; 95US-0384626.
XX PR 24-AUG-1998; 98US-0138775.
XX PA (HOFF ) ROCHE VITAMINS INC.
XX XX Erdenberger T, Hatch RT, Hermann T, Perkins JB;
XX PI Pero JG, Sloma A;
XX DR WPJ: 1999-418271/35.
XX DR P-PSDB; Y21801, Y21802, Y21803.
XX PT Use of bacterial strains for the over production of riboflavin
XX PS Example 2; Fig 3A-S; 61pp; English.
XX CC The invention relates to a method of producing riboflavin that comprises,
XX CC providing a recombinant bacterium containing an exogenously introduced
XX CC bacterial ORF (open reading frame) 3 (or ORF5) structural gene and an
XX CC exogenous promoter, culturing the bacterium, and recovery of the product.
XX CC The method is useful for the production of large quantities of riboflavin
XX CC (vitamin B2). The present sequence represents the complete nucleotide
XX CC sequence of B. subtilis rib operon (also see X81948 for the antisense
XX CC sequence).
XX SO Sequence 5567 BP; 1751 A; 1096 C; 1297 G; 1423 T; 0 other:

Query Match 11.7%; Score 120.4; DB 20; Length 5567;
Best Local Similarity 51.0%; Pred. No. 4.9e-19;
Matches 317; Conservative 0; Mismatches 296; Indels 9; Gaps 1;

```

```

Db 1296 gacattacgttacacacgcgcgtgcagccattacgagaaacccgcacgtgcagaa 1355
Qy 241 AAAATTATGATTGTGAAGATAGCAAAAGTAGTATGCCAACAAGA-----CAAT 291
Db 1356 ttgattacaaacttcgttatacaaaagtgctgcgcgcagagaagatcctaaccgctc 1415
Qy 292 TCGTTAGACACATGATGATGACGCTTACGGGCTCAGCGTATTGAGGTTGAATGCGTT 351
Db 1416 gtgcctggaagaaggatcagcatgtatgaagaagctcgtcattgagtgaaagagcctc 1475
Qy 352 GATGATGAACGGGCATCATCAATATACCAAGACTTTTAAAGCAAAAGCAACCACTG 411
Db 1476 ctgcagacccagcgcggaagcgtcgaatgaataattctgccttatagagacgcgctc 1535
Qy 412 CCACAAATTACAGTGAAGTATCTGCAAGTTAGTGTAAACAAGGAATGATATGA 471
Db 1536 ccgtacgcacgcctaaagcgcgtgcagcccttgacgcgcaagataagctacacagcggct 1595
Qy 472 CAAAGTCATGATTTACTATACAAAGAGTTAAACAAGATGCTATTAAGTTAAGCATCGA 531
Db 1596 gacagcaaatgcatcagcagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1655
Qy 532 CACGACGCGATGTTAACTGGAAGACGTACAGTTGATGATGCCAATATACTACA 591
Db 1656 caccaaagcattttagtcgcaggttcgcacagtgaaagcgcgaatccgcagcttaacctgc 1715
Qy 592 CGTATTCAGATGGAAGAAAACC 613
Db 1716 agactgcgcgaatgaacaaac 1737

RESULT 11
X81948/C
ID X81948 standard; DNA; 5567 BP.
XX AC X81948;
XX DT 10-SEP-1999 (first entry)
XX DE B. subtilis rib operon complementary nucleotide sequence.
XX KM Riboflavin; open reading frame; ORF; structural gene; promoter;
XX KM vitamin B2; Bacillus subtilis; rib operon; complementary; ds.
XX OS Bacillus subtilis.
XX FH Key
XX FH CDS
XX FT 1..5567
XX FT /tag= a
XX FT /note= "reading frame 1 translated protein (Y21804)"
XX FT 3..5567
XX FT /tag= b
XX FT /note= "reading frame 2 translated protein (Y21805)"
XX FT 2..5567
XX FT /tag= c
XX FT /note= "reading frame 3 translated protein (Y21806)"
XX PN US592538-A.
XX PD 20-JUL-1999.
XX PF 24-AUG-1998; 98US-0138775.
XX PR 11-SEP-1990; 90US-0581048.
XX PR 22-JUN-1989; 89US-0370378.
XX PR 21-APR-1992; 92US-0873572.
XX PR 06-FEB-1995; 95US-0384626.
XX PR 24-AUG-1998; 98US-0138775.
XX PA (HOFF ) ROCHE VITAMINS INC.
XX XX Erdenberger T, Hatch RT, Hermann T, Perkins JB;
XX PI Pero JG, Sloma A;

```

XX MPI: 1999-418271/35.
DR P-PSDB: Y21804, Y21805, Y21806.
XX
PT Use of bacterial strains for the over production of riboflavin
XX
PS Example 2; Fig 3A-S; 61pp; English.
XX
CC The invention relates to a method of producing riboflavin that comprises,
CC providing a recombinant bacterium containing an exogenously introduced
CC bacterial ORF (open reading frame) 3 (or ORF3) structural gene and an
CC exogenous promoter, culturing the bacterium, and recovery of the product.
CC The method is useful for the production of large quantities of riboflavin
CC (vitamin B2). The present sequence represents the complementary
CC nucleotide sequence of B. subtilis rib operon.
XX
SQ Sequence 5567 BP; 1422 A; 1296 C; 1097 G; 1752 T; 0 other;

Query Match 11.7%; Score 120.4; DB 20; Length 5567;
Best local Similarity 51.0%; Pred. No. 4,9e-19;
Matches 317; Conservative 0; Mismatches 296; Indels 9; Gaps 1;

QY 1 ATGATTATGCGATTCACTTGCATAATATGTTACAGGTCAACAGGTGTTATCCACC 60
DB 4452 ATGAACTGCGCTTAGATCTTGCAAGACAGGCGGAGACAGACCGAATCCATCCGCTC 4393

QY 61 GTTGCCCTTTGTAGTTAATGAAGTAGATTTGTGTATGTGTCACACTTGAGAAA 120
DB 4392 GTGCGCGCTTTGTGTTAAAGGACGACAAATGTCGAAATGGCGCCATTAAATAT 4333

QY 121 GGTGACAAAGATGGGAGGTTCAAGCACTGATGTCACAAACAAATGCTGAAGTGGC 180
DB 4332 GGTGAACCTATGACAGATTTCATCCATATGCTGTGAGACATGACAGAGGTGCC 4273

QY 181 ACGATTATATACGTTAGACCATGATGATTTGTTCAACACACCCCTGTGTAC 240
DB 4272 GACATTTACGTTACACTGCAACCGTCGACCATACGAAACACCGCATGTGAGAA 4213

QY 241 AAAATTATTATGTAGATAGCAAAAGTAGTATACCCACAAAGA-----CAAT 291
DB 4212 TTGATTATCAACTCTGTATCAAAAGAGTGTCTGTCGATGAGATCCTAATCCGCTT 4153

QY 292 TCGTTACACACATGTTGATGAGAGCTTACGGGCTACGCTATGAGTTGATGCGTT 351
DB 4152 GTGCTGGAAGAGGATCAGCATGATGAAGACCTGCGATTGAGTAAGGAGGCATC 4093

QY 352 GATGATGAACGGGATCACTATTATACCAAGACTTTTAAAGCAAAAGCAAGCACTG 411
DB 4092 CTGGCAGACAGGCGGAGAGGCTGAATGAATAATTTCTGCACCTTATAGAGACAGGCTT 4033

QY 412 CCACAAATTTACAGTAAGATCTGCAAGTTTATGATGTTAAACAGGAAATGATATGA 471
DB 4032 CCGTACGTCACGCTAAAGAGGCGCTGACGCTGACGCAAGATGCTACACAGCAGGT 3973

QY 472 CAAGTCAATGATTAATAAAGAGTTAAACAAGATGCTATAGTTAAGACATGA 531
DB 3972 GACGCAAAATGATCAGTCAGAGGTCGAAGAGATGCTCGCATACAGGAAACA 3913

QY 532 CACGACGCACTGTTAACTGAGACGTAAGTGAATGATGATGATATATCTACA 591
DB 3912 CACCAAAAGCATTTTAGTCGGAGTTGGCAGCAGTGAACCGACATCGAGCTTAACCTGC 3853

QY 592 CGTATTCAAGATGGAACCAAC 613
DB 3852 AGACTGCCGATGTAAACAAAC 3831

RESULT 12
V71600
ID V71600 standard; DNA; 5567 BP.
XX
AC V71600;

XX 01-MAR-1999 (first entry)
XX
XX DE Bacillus subtilis rib operon.
XX
XX KW Riboflavin; vitamin B2; riboflavin synthase; deaminase; reductase;
XX KM GTP cyclohydrolase; rib operon; ds.
XX
XX OS Bacillus subtilis strain 168.
XX
FH Key Location/Qualifiers
FT RBS 352..357
FT RBS /tag= a
FT CDS 364..681
FT /tag= b
FT /label= ORF6
FT 708..749
FT /tag= c
FT /note= "rho-independent transcription terminator"
FT 771..799
FT /tag= d
FT /label= PI
FT 1034..67
FT /tag= e
FT /note= "rho-independent transcription terminator"
FT RBS 1184..1195
FT /tag= f
FT 1422..1433
FT /tag= g
FT 1437..2186
FT /tag= h
FT /label= ORF5
FT 2178..2194
FT /tag= i
FT 2197..2844
FT /tag= j
FT /label= ORF4
FT 2528..2556
FT /tag= k
FT /label= ORF3
FT 4071..4085
FT /tag= n
FT 4088..4552
FT /tag= o
FT /product= riboflavin synthase beta subunit
FT 4545..4574
FT /tag= p
FT /label= P3
FT 4653..4659
FT /tag= q
FT 4665..5039
FT /tag= r
FT /label= ORF2
FT 5038..5090
FT /tag= s
FT /note= "rho-independent transcription terminator"
FT complement (5053..5567)
FT /tag= t
FT /label= ORF1

XX US5837528-A.
XX
XX PD 17-NOV-1998.
XX
XX PE 06-FEB-1995; 95US-0384626.
XX 11-SEP-1990; 90US-0581048.
PR 22-JUN-1989; 89US-0370378.

PR 21-APR-1992; 92US-0873572.
 PR 06-FEB-1995; 95US-0384626.
 XX
 PA (HOFF) HOFMANN LA ROCHE INC.
 XX
 PI Perkins JB, Pero JG, Sloma A;
 XX
 DR WPI: 1999-023456/02.
 DR P-PSDB: W81549-55.

PT Nucleic acid encoding proteins involved in bacterial synthesis of
 PT riboflavin - also mutant or recombinant bacteria producing
 PT riboflavin at high levels
 XX
 PS

Example 7; Fig 3A-S; 62pp; English.

CC This is the nucleotide sequence of the rib operon of *Bacillus*
 CC *subtilis*. The rib operon includes an open reading frame (ORF)
 CC encoding the beta-subunit (see W81553) of riboflavin synthase, and
 CC 6 other ORFs, of which ORF1 is a partial coding region encoding a
 CC 170-amino acid polypeptide (see W81555), ORF2 encodes a putative
 CC rib-specific reductase (see W81554), ORF3 encodes a putative GTP
 CC cyclohydrolase (see W81552), ORF4 encodes the riboflavin synthase
 CC alpha subunit (see W81551), ORF5 encodes a putative rib-specific
 CC deaminase (see W81550) and ORF6 encodes a 38 kDa protein (see
 CC W81549) of unknown function. The invention provides vectors and
 CC recombinant bacteria for overproducing riboflavin, in which nucleic
 CC acids overproducing riboflavin biosynthetic proteins are introduced
 CC into the chromosome of the host organism at multiple sites or in
 CC multiple copies. The rib operon, e.g. comprising ORFs 2-5 and
 CC the beta-riboflavin synthase gene, is used to make such recombinant
 CC bacteria. The rib operon can be deregulated by replacing its
 CC control regions with sequences that allow constitutive or
 CC unregulated expression. The recombinant bacteria produce at least
 CC 10 g/l riboflavin.
 XX
 SQ

Sequence 5567 BP: 1749 A; 1102 C; 1298 G; 1418 T; 0 other:

Query Match 11.7%; Score 120.4; DB 20; Length 5567;
 Best Local Similarity 51.0%; Pred. No. 4,9e-19;

Matches 317; Conservative 0; Mismatches 296; Indels 9; Gaps 1;

QY 1 ATGGATTATGCGATTCACTTGCATAATATGTACAGATCAAGAGTTTATCCACCC 60
 DB 1116 atgaagctgcttgaatcttcgaaagcgagagcgagacgacccatccacgcgc 1175
 QY 61 GTTGGCGTGTGTAGTATGATGAAGTAGATTGTTGTTGCTGCACACTTGAGAAA 120
 DB 1176 gtgcgagctgttcgtcgtlaaagagcaaatgtcgsaatgagcgccattaaatat 1235
 QY 121 GGTGACACATCGGAGGTTCAAGCACTTGATATGCGACACAAATAATCTGAAGTGG 180
 DB 1236 ggtlgaagccatcagcaagttcatgcatccatcgtcggagcaatgcagaaggtgcc 1295
 QY 181 ACGATTATATTACGTTAGTAGTATGATGATGTTGTTGTTCAACACCACTGTGTTAAC 240
 DB 1296 gacatttaactgaactgcgaacccgtcagccacttaagcaaaaacacgcctgtgcagaa 1355
 QY 241 AAAATTATGATTGTAAAGTAGCAAAAAGTAGTATACGACAAACAAGAAAA-----CAAT 291
 DB 1356 ttgatattcaactctgtatctaaagaagtgctgcgtgagtgagatctcaatccgctt 1415
 QY 292 TCGTTAGACACATGTTGATGACAGCTTACGGGCTTCACGGATTTTAAAGTTGAATGCGTT 351
 DB 1416 gtgcgtcgaagagggatcagcatgaaagagcgtgcgcatlgaaggaaagggatc 1475
 QY 352 GATGATGAACGGGATCACAATATATACCAAGACTTTTAAAGCAAAAGCAAAAGCAACTG 411
 DB 1476 ctgcgacagcagcgagagcgctgaatgaaatctctcactttatgaggaagcgctt 1535
 QY 412 CCACAATTACAGTGAAGATATCTGCAAGTTTATGATGTTAAACCAAGCAATGATTAATGGA 471

DB 1536 ccgtacgtcacgtataaagcgctgcagcccttcagcgcaagataagctaccagcagcggt 1595
 QY 472 CAAAGTCATGGATTACTTAACAAGAGGTAAACAGATGCTATTAAGTTAAGACATCGA 531
 DB 1596 gacagcaaatgatacagctcagagcgctgcgaagacagatgtctcgaatatacaggaataca 1655
 QY 532 CACGACGAGTGTAACTGAAGAGCTACAGTTGATGATCCCAATATACTACA 591
 DB 1656 caaccaagcattttagctgcgagttgcacagtgaaagcgcaacaatccgacttaaccctgc 1715
 QY 592 CGTATTCAGATGCAAAAACC 613
 DB 1716 agactgcggaatgatacaaac 1737

RESULT 13

ID 293752
 Z93752 standard; DNA: 5567 BP.

AC 293752;

XX 16-AUG-2000 (first entry)

DE Rib operon of *Bacillus subtilis*.

XX Riboflavin; rib operon; yeast; bacteria; transformation;

KW transcription element; gene expression; hair loss; skin inflammation;

KW growth disorder; vision; vision; sight; ds.

OS *Bacillus subtilis*.

FH Key Location/Qualifiers

FT CDS 1..5567

FT CDS 1..5567

FT CDS 1..5567

FT CDS 1..5567

FT CDS 1..5567

FT CDS 1..5567

FT CDS 1..5567

FT CDS 1..5567

FT CDS 1..5567

FT CDS 1..5567

FT CDS 1..5567

FT CDS 1..5567

FT CDS 1..5567

FT CDS 1..5567

FT CDS 1..5567

FT CDS 1..5567

FT CDS 1..5567

FT CDS 1..5567

FT CDS 1..5567

FT CDS 1..5567

FT CDS 1..5567

FT CDS 1..5567

```

FT      /product= Beta riboflavin synthase
FT      4545..4574
FT      /*tag= g
FT      /label= p3 promoter
FT      4653..4659
FT      RBS
FT      /*tag= r
FT      4665..5039
FT      CDS
FT      /*tag= s
FT      /label= ORF2
FT      5038..5090
FT      terminator
FT      /*tag= t
FT      /label= Rho independent transcription terminator
FT      complement (5053..5567)
FT      /*tag= u
FT      /label= ORF1
FT      EPI001026-A2.
PN      17-MAY-2000.
PD      22-JUN-1990: 99EP-0124888.
PE      22-JUN-1989: 89US-0370378.
PR      22-JUN-1990: 90EP-0111916.
XX      (HOFF ) HOFFMANN LA ROCHE & CO AG F.
XX      Perkins JB, Pero JG, Sloma A:
PI      WPI, 2000-320439/28.
XX      P-PSDB: Y83269, Y83270, Y83271.
DR      Novel vector, used for the large scale production of riboflavin,
PT      comprises a bacterial or yeast riboflavin biosynthetic protein under
XX      the control of a foreign transcription element
XX      Claim 1: Figure 3: 66pp: English.
XX      Vectors comprising a nucleic acid sequence of bacterial or
XX      yeast origin, coding for one or more riboflavin biosynthetic proteins,
XX      CC and one or more transcription elements not naturally associated with
XX      CC the nucleic acid sequence are new. The vector and bacteria transformed
XX      CC with it are used for the large scale production (over 10g/l) of
XX      CC riboflavin. The riboflavin produced can be used to treat disorders
XX      CC associated with riboflavin deficiency such as hair loss, skin
XX      CC inflammation, vision deterioration and growth failure.
XX      SO Sequence 5567 BP: 1751 A: 1097 C: 1296 G: 1423 T: 0 other:
Query Match 11.7%; Score 120.4; DB 21: Length 5567;
Best Local Similarity 51.0%; Pred. No. 4.9e-19;
Matches 317: Conservative 0; Mismatches 296; Indels 9; Gaps 1;
QY      1 ATGGATTATGCGATTGCAATATGTTACAGGTCAACAGGTGTTAATCCACCC 60
DB      1116 atgaagctgctgcttagatcttgcgaagcagggcgaagacagacccaatccgcgcc 1175
QY      61 GTTGGCGTGTGTAGTAAATGAAGTAGATTGTTGTTGTTGTCACACTTGAGAAAA 120
DB      11176 gtgcggcgctgtgtcgtlaaagcagcaaatgtcgcgaatggcgccattaaatat 1235
QY      121 GGTGACACGATCGGAGTTTCAAGCATTGATTCGCACACAAAATGCTGAAGTGGC 180
DB      1236 ggtgaagctcattgcagaatctcatgcacatcatgtcggagcaatgcagaggggtgcc 1295
QY      181 AGATTATATACGTAGTACGATGATGATTGTTGTTCAACACACCCCTGTAAAC 240
DB      1296 gacattacgttacacgcgaacccgctgcacattacgaagaaacacccgcctgtcaga 1355
QY      241 AAAATTATGATGTGAAGTAGCAAAAGTAGTATACGAACAAAGAAAA-----CAAT 291
DB      1356 ttgattatcaactctgtatcaaaagagctgtcgtgcgatgcagatcctaatccgctt 1415

```

```

QY      292 TCGTTAGACACATGTCGATGAGACGTTACGGCTCACGGTATTGAGGTTGAATGCCGT 351
DB      1416 gtgctgctggaagagggatcgcgcattgataaagaagctgtgcattgagtaaggaaagc 1475
QY      352 GATGATGAACGGGCATCACAAATTATACAGAGCTTTTAAAGCAAAAGCAAGCACTG 411
DB      1476 ctggcagaccagcggaagcgctgcatgaaanaattctcacttattgagcaagccctt 1535
QY      412 CCACAAATTACAGTGAAGTATCTGCAGTTTAGATGCTGAACCAAGCGAATATATGGA 471
DB      1536 ccgtacgtacccctaaagcggctgcacgccttgacgcggaagatagctccacgcgg 1595
QY      472 CAAAGTCATGATTAATTAACAAAGAGTTAAACAAAGATGCTATTAAGTTAAGACATCGA 531
DB      1596 gacagcaaatgatactcagctcgaagcctgcgaagacagatgctccagcaatcacgaa 1655
QY      532 CACGACGCGAGTGTAACTGGAAGACGTACAGTTGAATTAGATGATCCACAAATATCTACA 591
DB      1656 caccaaagcattttagtcggagttgacacagtgaaagccgacaatccgcgcttaacc 1715
QY      592 CGTATTCAAAGATGGAATAAAC 613
DB      1716 agactgcccgaatgtaacaaac 1737

RESULT 14
V52173/c
ID      V52173 standard; DNA: 14273 BP.
AC      V52173:
XX      23-OCT-1998 (first entry)
DE      Streptococcus pneumoniae genome fragment SEQ ID NO:40.
XX      Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;
KW      computer readable medium; vaccine; pharmaceutical composition; ds.
XX      Streptococcus pneumoniae.
OS      WO9818931-A2.
XX      PN 07-MAY-1998.
XX      PD 30-OCT-1997: 97WO-US19588.
XX      PE 31-OCT-1996: 96US-0029960.
XX      PR (HUMA-) HUMAN GENOME SCI INC.
XX      PA Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M;
PI      Kunsch CA, Rosen CA;
XX      DR WPI: 1998-272225/24.
XX      PT Computer-readable medium with recorded Streptococcus pneumoniae
PT      polynucleotide sequences - useful in diagnostic kits and assays, and
PT      pharmaceutical compositions and vaccines for Streptococcus
PT      pneumoniae
PS      Claim 1: Page 389-398; 1409pp: English.
XX      The present invention describes a computer readable medium which has
XX      CC the nucleotide sequences SEQ ID NO:1 to 391 (V52134 to V52524) recorded
XX      CC on it, or a representative fragment or a sequence at least 95% identical
XX      CC to SEQ ID NO: 1 to 391. The nucleotide sequences depicted in SEQ ID NO:1
XX      CC to 391 (V52134 to V52524) are genomic fragments from Streptococcus
XX      CC pneumoniae. The present invention also describes an isolated nucleic acid
XX      CC molecule encoding a homologue of any of the fragments of the S.pneumoniae
XX      CC genome (SEQ ID NO:1 to 391) where the nucleic acid molecule is produced
XX      CC by a process comprising: (a) screening a genomic DNA library using as a
XX      CC probe a target sequence defined by any of the sequences in SEQ ID NO:1

```

CC to 391, identifying members of the library which contain sequences
 CC that hybridise to the target sequence and isolating the nucleic acid
 CC molecules from the members; or (b) isolating mRNA, DNA or cDNA produced
 CC from an organism, amplifying nucleic acid molecules whose nucleotide
 CC sequence is homologous to amplification primers derived from the
 CC fragment of the S. pneumoniae genome to prime the amplification and
 CC isolating the amplified sequences. The computer readable medium can be
 CC used in a computer-based system for identifying fragments of the
 CC S. pneumoniae genome of commercial importance, or expression modulating
 CC fragments of the S. pneumoniae genome. Products from the present
 CC invention can be used in diagnosis kits and assays, and pharmaceutical
 CC compositions and vaccines for S. pneumoniae.

Sequence 14273 BP; 4357 A; 2428 C; 2694 G; 4794 T; 0 other:

Query Match 11.1%; Score 114.2; DB 19; Length 14273;
 Best Local Similarity 49.3%; Pred. No. 1.7e-17;
 Matches 387; Conservative 0; Mismatches 388; Indels 10; Gaps 3;

QY 1 ATGATATATCCGATTCACCTGCAATATGTGTAACAAGTCAACAGCTGTATCCACC 60
 DB 10928 ATGAAATATACCAATAAAGATGCGCAAAAGGGGCTGTACGTCATCCCAATCCCTATG 10869
 QY 61 GTTGGCGCTGTGTAGTAAAGTAGAGATTGTTGTTGTCGACACTTGAGAAA 120
 DB 10868 GTGGGCGCATTTATTTAAAGATATACATTTATGCGACAAGTTATCATGATTTT 10809
 QY 121 GGTGACAAAGCATCGGAGGTTCAAGCACTTGATATGCGACAACAACAAATGCTGAAGTGGC 180
 DB 10808 GGTGGCCCATCTGTAGAGAAATGCTTTAAACGTAGAGAAATCCCTGTGCGAGCG 10749
 QY 181 ACCATTTATTTAGTTAGTAGAGCCATAGTCATTTTGGTTCACACACCCCTGGTTAAC 240
 DB 10748 ACCCTTATATGACCTGTAACCTGTTGTACTGCGGAAACACCTCCCTGTATGAT 10689
 QY 241 AAATTTATTTATGTAAGTAGCAAGCAAGTATACGCAACAAAGA-----CAATT 292
 DB 10688 GGTATATCATAGTATGATATACAGAGTAGTCTATTGGAAGCCATGATATCCATT 10629
 QY 293 CGTTAGACACACATGTTGATGAGACGTTACGGGCTCAGCGTATGAGGTGAATGCGTTG 352
 DB 10628 GTATCTGGAAGAGGATTAAGATCTTGAGGAAATATCTTCAAGTTACTGTTGGAATT 10569
 QY 353 ATGATGACAGCGCATCATTAATACCAACACTTTTAAAGCA-AAAAGAACCACTG 411
 DB 10568 TTGAAATATAGTGTCTTAATCTTAATAAAAGTTTGAAGATATATTACCCAGCATGTA 10509
 QY 412 CCACAATTTACAGTGAAGATCTGCAAGTTTGAATGTTAAACAAGCGAATGATATGA 471
 DB 10508 CCTATGTTTTTATGAATAATGCAATGCAATGATGGAATAAATAGCCACTAAACAAT 10449
 QY 472 CCAAGTCATGATTAATTAACAAGAGGTTAAACAAGATGCTATTAAGTTAAGACATGA 531
 DB 10448 CAATCCAAATGATTAATGAAAGAACGAAACAAACATGTCATGTAACGACACCAT 10389
 QY 532 CAGAGCCATGTTAATGGAAGACGACAGTTAGTGAATGATGCAACATATNACTACA 591
 DB 10388 GTTAGTGCAATTTAGTGGGAGTCAATCTGTTATTCAGACGATCCCTTGGTCGACATGT 10329
 QY 592 CGATTTCAAGTGAAGAAACCCATATAAGTATATGTTAAGTCTGGGAATATTCAT 651
 DB 10328 AGATTTGAGAGAGGAAATATCCATATCCGATATATGCAATACATTTAAGCAAGCTCT 10269
 QY 652 TTTA-ATCAGCAATTTATCAAGATGAATCAACACCAATTTGATATATCTGAAAAATCC 710
 DB 10268 CTTACCTCTAAATCTGAAAAACAGCAATGATTAATAAACTTACATTTGCCACTCTCT 10209
 QY 711 AAATTTAACAAGCATTAACACACATATTTGAATATTTACTTGAAGTCTGTGATTTAAC 770
 DB 10208 GAAGACAAAAATAAATGAAGCTATATCAAAATCATGGCTGTGAAATCTTTCATTAAG 10149
 QY 771 AACAA 775

DB 10148 AAAAA 10144

RESULT 15
 ID X77179 standard; DNA; 1101 BP.
 XX X77179;
 AC X77179;

04-AUG-1999 (first entry)

S. pneumoniae ribg polypeptide encoding DNA.

Streptococcus pneumoniae: pyrimidine deaminase; pyrimidine reductase;
 ribg; immune response; antimicrobial: H. pylori infection; cancer;
 gastrointestinal carcinoma; gastric ulcer; gastritis; ss.
 Streptococcus pneumoniae.

MO9927126-A1.

03-JUN-1999.

23-NOV-1998: 98WO-US25010.

25-NOV-1997: 97US-0979616.

(SMK) SMITHKLINE BEECHAM CORP.

Black MT, Depouck C, Fedon JC, Hodgson JE, Jaworski D;
 Knowles DJC, Kosmatka AL, Lonetto MA, Mooney J, Nicholas RO;
 Palmer LM, Shilling LK, Stodola RK, Wang M, Warren RL;
 zhong Y;

WPI: 1999-347726/29.

P-PSDB: Y17949.

New Streptococcus pneumoniae ribg polypeptide and polynucleotide
 useful in the treatment of gastric ulcer and gastritis

Claim 6; Page 5; 48pp; English.

The invention relates to Streptococcus pneumoniae pyrimidine deaminase
 and pyrimidine reductase (ribg). S. pneumoniae ribg and its antagonists a
 re used to treat individuals in need of the polypeptide. Disease related
 to expression or activity of ribg can be determined by analysing the
 nucleic acid sequence encoding ribg or detecting the ribg polypeptide in
 a sample. ribg can also be used to identify antagonists or agonists.
 ribg, or its related nucleic acid, also has use as a vaccine to induce
 an immunological response in an animal. Antimicrobial compounds (e.g.
 agonists and antagonists of ribg), especially broad-spectrum antibiotics,
 may be of use in the treatment of Helicobacter pylori infection. This
 should decrease the advent of H. pylori-induced cancers, such as
 gastrointestinal carcinoma. The treatment should also cure gastric ulcers
 and gastritis. The present sequence represents a DNA encoding the
 S. pneumoniae ribg polypeptide.

Sequence 1101 BP; 392 A; 180 C; 216 G; 313 T; 0 other:

Query Match 10.9%; Score 112.6; DB 20; Length 1101;
 Best Local Similarity 49.2%; Pred. No. 2.1e-17;
 Matches 386; Conservative 0; Mismatches 389; Indels 10; Gaps 3;

QY 1 ATGATATATGCGATTCACCTGCAATATGTAACAAGTCAACAGCTGTATATCCACC 60
 DB 19 atgaattagcaataaactgacaaaaggggctgtacgtaacatccatcctatg 78
 QY 61 GTTGGCGCTGTGTAGTAAAGTAGAGTTGTTGATTTGGTGCACACTTGAGAAAA 120
 DB 79 gttygcgaattatgttaaagataatcacatcatcgcgaaggttatcatgatgttttt 138


```
OY 121 GGTGACAGCATGGAGGCTTCAAGCACTTGATATGGCACAAACAAATGCTGAGGTGCG 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 139 gggggcccaacatgctgagagaataatgctcttaaaacgtgagaaatccctgctggagcg 198
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 181 ACGATTTATATTACGTTAGAGCCATGTAGTCAATTTGGTTCAACACCACTGTGTTAAC 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 199 acgcttatgtaaacctgaaacctgtgtcacttcggaacaacctccctgatatagat 258
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 241 AAATTTGTGATGTAGATAGCAAAAGTAGTATACCAACAAAGA-----CAATT 292
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 259 gctataatcgtatgtagtatacagaagtagtcaatggaagcctagactgtaacctatc 318
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 293 CGTTAGACACATGATGATGAGACGTTACGGGCTCACGGTATTGAGTTGATGCGTTG 352
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 319 gtatctgaaaaagagaaagatacttgaggaataatacttcaagtlactgttgaatt 378
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 353 ATGATGAACGGGCATCACAATTATACCAAGACTTTTAAAGCA-AAAGCAAGCAACTG 411
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 379 ttaaaaaatgagtgcttaacttaataaaagtttagaaaglatattaccagcatgta 438
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 412 CCACAAATTACAGTGAAGTATCTGCAAGTTTATGATGTAACACGCAATGATTAATGA 471
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 439 ccctatgttlttgaatatgtcaatgtcgaatggaataatagccactaaacaat 498
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 472 CAAAGTCAATGATTTACTTAACAAGAGGTTTAAACAGATGCTATAAGTTTACATCGA 531
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 499 caatccaaatggaatgaagaagaagaagaagaagcatgtgcacgtttaggaactat 558
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 532 CACGACGCAAGTGTAACTGGAAGACGTACAGTTGAATTAGATGCCAATATATCTACA 591
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 559 gttagtcaattatgtgaggaagtcacatcgttatacaagacgtcccttgcgacatgt 618
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 592 CGTATTCAAGATGCAAAAACCCCTATATAAGTATATTGTCTAGTCTGGAAATATTCAT 651
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 619 agattggaagaaggaataacccatccgtatcatatcgaatcaacatllaagaacctcc 678
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 652 TTTA-ATCAGCAAAATTATCAGATGAATCAACACAATTTGATATATCTGAAAAATCC 710
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 679 cttaaccttaaaatcgtaaaaaacgcaaatgatacttaaaacttaccatgcactccctc 738
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 711 AAATTTAACAAGCAATCAACACATATTTGAATTTACTTGAAGTCTGTGATTTAAC 770
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 739 gaagacaaaaataaaatgaagctatcacaatcatgctgtgaatacttccataaag 798
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 771 AACAA 775
    |||||
Db 799 aaaaa 803
    |||||
```

Search completed: November 19, 2000, 05:07:08
Job time: 6894 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 19, 2000, 02:32:43 ; Search time 111.68 Seconds
(without alignments)
1393.512 Million cell updates/sec

Title: US-08-978-456-1

Perfect score: 1029
Sequence: 1 ATGCATTATGCATTCACT.....TACTTTCAGAAAGAGTGA 1029

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 262060 seqs, 75620727 residues

Total number of hits satisfying chosen parameters: 524120

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/2/ina/5C_COMB.seq: *
3: /cgn2_6/prodata/2/ina/5D_COMB.seq: *
4: /cgn2_6/prodata/2/ina/5D_COMB.seq: *
5: /cgn2_6/prodata/2/ina/6_COMB.seq: *
6: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq: *
7: /cgn2_6/prodata/2/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1029	100.0	1029	US-08-978-456-1	Sequence 1, Appl
2	1029	100.0	1029	US-08-978-456-1	Sequence 1, Appl
3	1029	100.0	1029	US-08-978-456-1	Sequence 1, Appl
4	352.2	34.2	1269	US-08-978-456-3	Sequence 3, Appl
5	112.6	10.9	3417	US-08-978-458-7	Sequence 7, Appl
6	112.6	10.9	3417	US-08-978-458-7	Sequence 7, Appl
7	76.2	7.4	1230	US-08-741-327E-14	Sequence 14, Appl
8	55.6	5.4	7218	US-08-232-463-14	Sequence 14, Appl
9	51.2	5.0	3095	5231168-1	Patent No. 5231168
10	47.8	4.6	9636	US-08-323-1708-1	Sequence 1, Appl
11	46.4	4.5	5589	US-08-465-795-1	Sequence 1, Appl
12	43.4	4.2	19124	US-08-487-8268-13	Sequence 13, Appl
13	40.4	4.0	4766	PCT-US93-07261-10	Sequence 10, Appl
14	40.4	3.9	665	US-08-883-795A-36	Sequence 36, Appl
15	39.2	3.8	4430	US-08-918-914-2	Sequence 2, Appl
16	38.8	3.8	1186	US-08-731-722-5	Sequence 5, Appl
17	38.8	3.8	3781	US-08-688-988-5	Sequence 5, Appl
18	38.6	3.8	19124	US-08-487-8268-13	Sequence 13, Appl
19	38.6	3.8	2817	PCT-US93-05944-1	Sequence 1, Appl
20	38.2	3.7	2530	US-08-450-351-1	Sequence 1, Appl
21	38.2	3.7	2520	US-08-450-351-3	Sequence 3, Appl
22	38.2	3.7	4248	US-08-678-614-1	Sequence 1, Appl
23	38	3.7	1939	US-07-715-751B-2	Sequence 2, Appl
24	37.8	3.7	10968	US-08-680-327-2	Sequence 2, Appl
25	37.8	3.7	2058	US-08-749-391-1	Sequence 1, Appl
26	37.2	3.6	1258	US-08-335-518-1	Sequence 1, Appl

27	37	3.6	2237	2	US-08-487-135B-1	Sequence 1, Appl
28	37	3.6	2237	3	US-08-915-972A-1	Sequence 1, Appl
29	37	3.6	2237	4	US-09-177-909-1	Sequence 1, Appl
30	36.6	3.6	446	5	US-09-097-541-1	Sequence 1, Appl
31	36.6	3.6	3926	4	US-08-731-722-1	Sequence 1, Appl
32	36.6	3.6	3926	4	US-08-731-722-1	Sequence 1, Appl
33	36.6	3.6	8920	4	US-08-446-855A-1	Sequence 1, Appl
34	36.6	3.6	56516	4	US-08-996-306-1	Sequence 1, Appl
35	36.4	3.5	3850	5	US-08-329-709-34	Sequence 34, Appl
36	36	3.5	3659	1	US-08-220-151-72	Sequence 72, Appl
37	36	3.5	3659	1	US-08-303-124-12	Sequence 12, Appl
38	36	3.5	3659	1	US-08-413-118-72	Sequence 72, Appl
39	36	3.5	3659	2	US-08-224-657-48	Sequence 48, Appl
40	36	3.5	3659	3	US-08-184-009-77	Sequence 77, Appl
41	36	3.5	3659	3	US-08-417-210A-67	Sequence 67, Appl
42	36	3.5	3659	3	US-08-480-697B-12	Sequence 12, Appl
43	36	3.5	3659	4	US-08-458-356-77	Sequence 77, Appl
44	36	3.5	3659	5	US-08-473-446-72	Sequence 72, Appl
45	36	3.5	3660	2	US-08-475-063-25	Sequence 25, Appl

ALIGNMENTS

RESULT 1
US-08-978-456-1
Sequence 1, Application US/08978456
Patent No. 6010881
GENERAL INFORMATION:
APPLICANT: Palmer, Leslie M.
TITLE OF INVENTION: No. 6010881el r1bG
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESS: Dechert Price & Rhoads
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
City: Philadelphia
STATE: PA
COUNTRY: US
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,456
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dickinson, Todd O
REGISTRATION NUMBER: 28,354
REFERENCE/DOCKET NUMBER: P50444-9
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-994-2252
TELEFAX: 215-994-2222
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1029 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-978-456-1
Query Match 100.0%; Score 1029; DB 5; Length 1029;
Best Local Similarity 100.0%; Pred. No. 2e-246;
Matches 1029; Conservative 0; Mismatches 0; Gaps 0;
OY 1 ATGCATTATGCATTCACTGCAATATGTCAGAGGTCAACAGGTGTTATCCACCC 60

```

|||||
Db 1 ATGATTTATGCATTCACACTGCAATATATGTAACAAGTCAAAACAGGTATATCCACCC 60
OY 61 GTTGGCGCTGTTGATTAATGAGGTAGAGTTGTTGATTTGCTCACACTTGAGAAAA 120
Db 61 GTTGGCGCTGTTGATTAATGAGGTAGAGTTGTTGATTTGCTCACACTTGAGAAAA 120
OY 121 GGTGACAAAGCATGCGAGAGTTCAAGCAGTTGATGCGACAAACAAATGCTGAAGTGG 180
Db 121 GGTGACAAAGCATGCGAGAGTTCAAGCAGTTGATGCGACAAACAAATGCTGAAGTGG 180
OY 181 ACGATTTATATTAATGAGTTAGAGCCATGTAATGTTGTTCAACACACCTGTTGTAAC 240
Db 181 ACGATTTATATTAATGAGTTAGAGCCATGTAATGTTGTTCAACACACCTGTTGTAAC 240
OY 241 AAAATTTATGATTTGTAAGTATAGCAAAAGTAGTATAGCAACAAACAAATTCGTTAGAC 300
Db 241 AAAATTTATGATTTGTAAGTATAGCAAAAGTAGTATAGCAACAAACAAATTCGTTAGAC 300
OY 301 ACACATGTTGATGAGAGCTTACGCGCTCACGATGAGTTGATGCGTTGATGATGA 360
Db 301 ACACATGTTGATGAGAGCTTACGCGCTCACGATGAGTTGATGCGTTGATGATGA 360
OY 361 CGGCGATCACAATTTATACCAAGACTTTTAAAGCAAAAGCAAGCACTGCCCAAT 420
Db 361 CGGCGATCACAATTTATACCAAGACTTTTAAAGCAAAAGCAAGCACTGCCCAAT 420
OY 421 AAGAGTCAAGTATCTGCAGCTTATAGTGTGTAACCAAGCAATGATTAATGACAAACTCA 480
Db 421 AAGAGTCAAGTATCTGCAGCTTATAGTGTGTAACCAAGCAATGATTAATGACAAACTCA 480
OY 481 TCGATTTACTAAGAAAGAGTTAAACAGATGCTATAGTTAAGATGAGACGACGCA 540
Db 481 TCGATTTACTAAGAAAGAGTTAAACAGATGCTATAGTTAAGATGAGACGACGCA 540
OY 541 GTGTTAACTGGAAGAGCTGAGTTGATTAAGATGATCACAATATCTACAGCTATTC 600
Db 541 GTGTTAACTGGAAGAGCTGAGTTGATTAAGATGATCACAATATCTACAGCTATTC 600
OY 601 GATGAAAAAACCCTTAAAGATATATGCTGTAAGTGTGGAATATTTATTCAG 660
Db 601 GATGAAAAAACCCTTAAAGATATATGCTGTAAGTGTGGAATATTTATTCAG 660
OY 661 CAAATTTATCAAGATGATCAACACCAATTTGATATATGTAATCAAAATTTACA 720
Db 661 CAAATTTATCAAGATGATCAACACCAATTTGATATATGTAATCAAAATTTACA 720
OY 721 ACGAATCAAAACATATGAAATTTATTTACTGGAAGCTTGTGATTTAAACAATTC 780
Db 721 ACGAATCAAAACATATGAAATTTATTTACTGGAAGCTTGTGATTTAAACAATTC 780
OY 781 CACATTTTATTAAGAGAGAGTTGGAAGTTGCTAGTGAAGCAGGTCACCACTACT 840
Db 781 CACATTTTATTAAGAGAGAGTTGGAAGTTGCTAGTGAAGCAGGTCACCACTACT 840
OY 841 TAGAATTTCCATCTATATATAGATGAATTTATTTCTATATTAATCCCGAATTA 900
Db 841 TAGAATTTCCATCTATATATAGATGAATTTATTTCTATATTAATCCCGAATTA 900
OY 901 GCGGATCTGGAATTTATCAATTTTATCAACAAATGATGATGATGATGATGATG 960
Db 901 GCGGATCTGGAATTTATCAATTTTATCAACAAATGATGATGATGATGATGATG 960
OY 961 AACCAATTTGAAATGTTGATTCGAGTTATTAATCAAAATGTTAAATTAATTTAC 1020
Db 961 AACCAATTTGAAATGTTGATTCGAGTTATTAATCAAAATGTTAAATTAATTTAC 1020
OY 1021 AAGAGTGA 1029
Db 1021 AAGAGTGA 1029

```

RESULT 2

```

US-08-977-554-7
; Sequence 7, Application US/08977554
; Patent No. 5891672
; GENERAL INFORMATION:
; APPLICANT: Palmer, Leslie M.
; APPLICANT: Fedon, Jason C.
; APPLICANT: Warren, Richard L.
; APPLICANT: Traini, Christopher M.
; APPLICANT: Wang, Min
; APPLICANT: Jaworski, Deborah D.
; APPLICANT: Mooney, Jeffrey
; APPLICANT: Debouck, Christine
; APPLICANT: Zhong, Yixi
; APPLICANT: Black, Michael
; TITLE OF INVENTION: riba
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/977,554
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/02318
; FILING DATE: 19-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, O. Todd
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: P50444-07
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215/994-2252
; TELEFAX: 215/994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3336 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-977-554-7

Query Match 100.0%; Score 1029; DB 3; Length 3336;
Best Local Similarity 100.0%; Pred. No. 2.9e-246;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ATGATTTATGCATTCACACTGCAATATATGTAACAAGTCAAAACAGGTATATCCACCC 60
Db 1 ATGATTTATGCATTCACACTGCAATATATGTAACAAGTCAAAACAGGTATATCCACCC 60
OY 61 GTTGGCGCTGTTGATTAATGAGGTAGAGTTGTTGATTTGCTCACACTTGAGAAAA 120
Db 61 GTTGGCGCTGTTGATTAATGAGGTAGAGTTGTTGATTTGCTCACACTTGAGAAAA 120
OY 121 GGTGACAAAGCATGCGAGAGTTCAAGCAGTTGATGCGACAAACAAATGCTGAAGTGG 180
Db 121 GGTGACAAAGCATGCGAGAGTTCAAGCAGTTGATGCGACAAACAAATGCTGAAGTGG 180
OY 181 ACGATTTATATTAATGAGTTAGAGCCATGTAATGTTGTTCAACACACCTGTTGTAAC 240
Db 181 ACGATTTATATTAATGAGTTAGAGCCATGTAATGTTGTTCAACACACCTGTTGTAAC 240
OY 241 AAAATTTATGATTTGTAAGTATAGCAAAAGTAGTATAGCAACAAACAAATTCGTTAGAC 300

```

```

Db 241 AAAATTATGTTGTAGATAGCAAAGATGATACGACACAAAAGACAAATTCCGTAGAC 300
|
|
|
QY 301 ACACATGGTATGAGACGTTACGGGCTCACGGTATTGAGGTTGAATGCGTTGATGATGAA 360
|
|
|
Db 301 ACACATGGTATGAGACGTTACGGGCTCACGGTATTGAGGTTGAATGCGTTGATGATGAA 360
|
|
|
QY 361 CGGGCATCACAATTATACCAAGACTTTTTTAAAGCAAAAGCAACCACTGCCACAATT 420
|
|
|
Db 361 CGGGCATCACAATTATACCAAGACTTTTTTAAAGCAAAAGCAACCACTGCCACAATT 420
|
|
|
QY 421 ACAGTGAAGATCTCGCAACTTTAGATGGTAAACCAAGCAATGATATAGGACAAAGTCAA 480
|
|
|
Db 421 ACAGTGAAGATCTCGCAACTTTAGATGGTAAACCAAGCAATGATATAGGACAAAGTCAA 480
|
|
|
QY 481 TGGATTACTAACAAAGAGGTTAAACAAGATGCTATATAGTTAAGACATCGACAGACGA 540
|
|
|
Db 481 TGGATTACTAACAAAGAGGTTAAACAAGATGCTATATAGTTAAGACATCGACAGACGA 540
|
|
|
QY 541 GTGTTAACTGGAAGACGTACAGTTGAATTATGATGATCCCAATTATCTACACGATTCGA 600
|
|
|
Db 541 GTGTTAACTGGAAGACGTACAGTTGAATTATGATGATCCCAATTATCTACACGATTCGA 600
|
|
|
QY 601 GATGAAAAAACCCCTATAAAAGTAAATTTGCTTAAGTCTGGGAATATTCATTTATATCAG 660
|
|
|
Db 601 GATGAAAAAACCCCTATAAAAGTAAATTTGCTTAAGTCTGGGAATATTCATTTATATCAG 660
|
|
|
QY 661 CAAATTTATCAGATGATCAACACCACCAATTTGGATATATCTGAAATCCAAATTTAACA 720
|
|
|
Db 661 CAAATTTATCAGATGATCAACACCACCAATTTGGATATATCTGAAATCCAAATTTAACA 720
|
|
|
QY 721 AGCAATCAAAACACATATTTGAAATTTATTTACTTGAACTCTTGATTTAACAACATCTT 780
|
|
|
Db 721 AGCAATCAAAACACATATTTGAAATTTATTTACTTGAACTCTTGATTTAACAACATCTT 780
|
|
|
QY 781 CACATTTATATATAAAGAGAGGTTGGACGTTGCTAGTGGAGGACAGTCCAAACACTACT 840
|
|
|
Db 781 CACATTTATATATAAAGAGAGGTTGGACGTTGCTAGTGGAGGACAGTCCAAACACTACT 840
|
|
|
QY 841 TCAGAAATTCATCATATTTATATATAGATGAAATTTATTTCTCTATTATGCCCGAAATTAAT 900
|
|
|
Db 841 TCAGAAATTCATCATATTTATATATAGATGAAATTTATTTCTCTATTATGCCCGAAATTAAT 900
|
|
|
QY 901 GCGCGATCTGGAATTTATCAATTTTATCAAAACAATGATGATGATTGAGATACCGATGCG 960
|
|
|
Db 901 GCGCGATCTGGAATTTATCAATTTTATCAAAACAATGATGATGATTGAGATACCGATGCG 960
|
|
|
QY 961 AACCAATTTGAAATTTGTTCAATTCGCGATTATTAACCAAAATGTTAAATTAACCTTAGCA 1020
|
|
|
Db 961 AACCAATTTGAAATTTGTTCAATTCGCGATTATTAACCAAAATGTTAAATTAACCTTAGCA 1020
|
|
|
QY 1021 AAGAAGTGA 1029
|
|
|
Db 1021 AAGAAGTGA 1029
|
|
|
RESULT 3
US-08-978-456-7
; Sequence 7, Application US/08978456
; Patent No. 6010881
; GENERAL INFORMATION:
; APPLICANT: Palmer, Leslie M.
; TITLE OF INVENTION: No. 6010881el r1bG
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM type: Diskette
;

```

```

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTED for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978.456
CLASSIFICATION: 435
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dickinson, Todd O
REGISTRATION NUMBER: 28,354
REFERENCE/DOCKET NUMBER: P50444-9
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-994-2252
TELEFAX: 215-994-2222
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3336 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-978-456-7

Query Match      100.0%; Score 1029; DB 5; Length 3336;
Best Local Similarity 100.0%; Pred. No. 2,9e-246;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGATTTATGCGATTCACACTTGCAGATATGCTACAAAGTCACAAAGCGTGTATATCCACC 60
Db 1 ATGATTTATGCGATTCACACTTGCAGATATGCTACAAAGTCACAAAGCGTGTATATCCACC 60

61 GTTGCGCGTGTGTAGTAAATGAAGTGAAGATTGTTGGATTGGTGCACACTTGAAGAAA 120
Db 61 GTTGCGCGTGTGTAGTAAATGAAGTGAAGATTGTTGGATTGGTGCACACTTGAAGAAA 120

121 GGTGCAAGCATGCGGAGGTTCAACGACTTGATATGCGCAACAAATCTGAAGTGGC 180
Db 121 GGTGCAAGCATGCGGAGGTTCAACGACTTGATATGCGCAACAAATCTGAAGTGGC 180

181 ACGATTATATTACGTTAGAGCATGTAGTCAATTTTGGTTCACACACCCTGTGTATAC 240
Db 181 ACGATTATATTACGTTAGAGCATGTAGTCAATTTTGGTTCACACACCCTGTGTATAC 240

241 AAAATTATGATTGTTAAGATAGCAAAAGTAGTACGCCACAAAAGACAAATTCGTTAGC 300
Db 241 AAAATTATGATTGTTAAGATAGCAAAAGTAGTACGCCACAAAAGACAAATTCGTTAGC 300

301 ACACATGGTGTAGAGCGTTACGGGCTCAGCGTATTGAGGTTGAATGCGTTGATGTGAA 360
Db 301 ACACATGGTGTAGAGCGTTACGGGCTCAGCGTATTGAGGTTGAATGCGTTGATGTGAA 360

361 CGGCGCATCACAATTATACCAAGACTTTTAAAGCAAAAAGCAAAACCACTGCCACAAAT 420
Db 361 CGGCGCATCACAATTATACCAAGACTTTTAAAGCAAAAAGCAAAACCACTGCCACAAAT 420

421 ACAGTGAAGATATCGCAAGTTAGATGTTAAACAAGGAATGATTAATGGAACAAAGTCAA 480
Db 421 ACAGTGAAGATATCGCAAGTTAGATGTTAAACAAGGAATGATTAATGGAACAAAGTCAA 480

481 TGGATTACTTAACAAAGAGGTTAAACAGAGTGTCTAAGTTAAAGATCGACGACGACGA 540
Db 481 TGGATTACTTAACAAAGAGGTTAAACAGAGTGTCTAAGTTAAAGATCGACGACGACGA 540

541 GTGTTAACTGGAAGCGTACAGTGAATTAAGTATGATCCCAATATATCTACAGTATTCAA 600
Db 541 GTGTTAACTGGAAGCGTACAGTGAATTAAGTATGATCCCAATATATCTACAGTATTCAA 600

601 GATGGAAGAAACCCATTAAGTATATTTGCTAGTCTGGGAATATTCATTTATATCG 660

```

Db 601 GATGAAAAAACCCTATAAAGTAATATGCTGAAGCTGGGAATATTCATTTTAATCAG 660
QY 661 CAAATTTATGATGATCAACACCAATTTGATATATGTAAGAAATCCAAATTTAACA 720
Db 661 CAAATTTATCAAGATGATCAACACCAATTTGATATATGTAAGAAATCCAAATTTAACA 720
QY 721 AGCAATCAACACATATTTGAAATTTTACTTGAAGCTTGTGATTTAACAACAAATCTT 780
Db 721 AGCAATCAACACATATTTGAAATTTTACTTGAAGCTTGTGATTTAACAACAAATCTT 780
QY 781 CACATTTATATAAAGAGAGCTTGGAACTTGTGATGAGGCGGTCCAACTACT 840
Db 781 CACATTTATATAAAGAGAGCTTGGAACTTGTGATGAGGCGGTCCAACTACT 840
QY 841 TCAGATCTCCATCTATATATATGATGAAATTTATCTCTATATGAGCCCGAAATTAAT 900
Db 841 TCAGATCTCCATCTATATATATGATGAAATTTATCTCTATATGAGCCCGAAATTAAT 900
QY 901 GGGGATCTGGAATTTATCAATTTTATCAAAACAAATGATGATGAGATACAGATGCG 960
Db 901 GGGGATCTGGAATTTATCAATTTTATCAAAACAAATGATGATGAGATACAGATGCG 960
QY 961 AACCAATTTGAAATTTGATTCATTCGAGTTATTAATCAAAATGTTAAATTAATTAAG 1020
Db 961 AACCAATTTGAAATTTGATTCATTCGAGTTATTAATCAAAATGTTAAATTAATTAAG 1020
QY 1021 AAGAAGTGA 1029
Db 1021 AAGAAGTGA 1029

RESULT 4

US-08-978-456-3

: Sequence 3, Application US/08978456

: Patent No. 6010881

: GENERAL INFORMATION:

: APPLICANT: Palmer, Leslie M.

: TITLE OF INVENTION: No. 6010881el ribg

: NUMBER OF SEQUENCES: 7

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: Dechert Price & Rhoads

: STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre

: CITY: Philadelphia

: STATE: PA

: COUNTRY: US

: ZIP: 19103

: COMPUTER READABLE FORM:

: MEDIUM TYPE: Diskette

: COMPUTER: IBM Compatible

: OPERATING SYSTEM: DOS

: SOFTWARE: FastSeq for Windows Version 2.0

: CURRENT APPLICATION DATA:

: APPLICATION NUMBER: US/08/978,456

: FILING DATE:

: CLASSIFICATION: 435

: PRIOR APPLICATION DATA:

: APPLICATION NUMBER:

: FILING DATE:

: ATTORNEY/AGENT INFORMATION:

: NAME: Dickinson, Todd Q

: REGISTRATION NUMBER: 28,354

: REFERENCE/DOCKET NUMBER: P5044-9

: TELECOMMUNICATION INFORMATION:

: TELEPHONE: 215-994-2252

: TELEFAX: 215-994-2222

: TELETYPE:

: INFORMATION FOR SEQ ID NO: 3:

: SEQUENCE CHARACTERISTICS:

: LENGTH: 1269 base pairs

: TYPE: nucleic acid

: STRANDEDNESS: double

: TOPOLOGY: linear

: US-08-978-456-3

Query Match 34.2%; Score 352.2; DB 5; Length 1269;
Best Local Similarity 88.4%; Pred. No. 1,1e-78;
Matches 473; Conservative 0; Mismatches 44; Indels 18; Gaps 8;

QY 1 ATGATTAATGCGATTCACTTGCATAATATGATACAAAGTCAAAAGGTGTTAATCCACC 60
Db 734 ATGATTAATGCGATTCACTTGCATAATATGATACAAAGTCAAAAGGTGTTAATCCACC 793
QY 61 GTTGCGCTGTGTAGTAAATGAAGTAGAGATTTGTGATTTGTTGATGACACTGAGAAA 120
Db 794 GTTGCGCTGTGTAGTAAATGAAGTAGAGATTTGTGATTTGTTGATGACACTGAGAAA 853
QY 121 GGTGACAGCATGGGAGGTTCACAGCACTTGATATGACACAAACAAATGCTGAAGTGGC 180
Db 854 GGTGACAGCATGGGAGGTTCACAGCACTTGATATGACACAAACAAATGCTGAAGTGGC 913
QY 181 ACGATTTATATACGTTAGAGCCATGTATCTATTTGTTCAACACCACTGTGTTAAC 240
Db 914 ACGATTTATATACGTTAGAGCCATGTATCTATTTGTTCAACACCACTGTGTTAAC 973
QY 241 AAATTTATGTTTGAATATGCAAAAGTAGTA-TACGCAACAAACAAAT-CTTAG 298
Db 974 AAATTTATGTTTGAATATGCAAAAGTAGTA-TACGCAACAAACAAAT-CTTAG 1033
QY 299 ACACACAT-GGTGATGAGACGTTACGGCTCACGGTATGAGTTGATCGTT----- 351
Db 1034 ACACACATGGGTGATGAGACGTTACGGCTCACGGTATGAGTTGATCGTT----- 1093
QY 352 GATGATGACGGGCGATCACTTATACCAAGACTTTT---AAAGCAAAAGCAACAA 408
Db 1094 GATGATGACGGGCGATCACTTATACCAAGACTTTT---AAAGCAAAAGCAACAA 1153
QY 409 CTGGCAC--AAATACAGTGAAGTATC-TGCAAGTTTACATGTTAACAAGG---AAT 462
Db 1154 CTGGCACAAATTTACAGTGAAGTATC-TGCAAGTTTACATGTTAACAAGG---AAT 1213
QY 463 GATATGACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 517
Db 1214 GATATGACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1268

RESULT 5

US-08-978-458-7

: Sequence 7, Application US/08978458

: Patent No. 5932701

: GENERAL INFORMATION:

: APPLICANT: Palmer, Leslie M.

: APPLICANT: Fedon, Jason C.

: APPLICANT: Warren, Richard L.

: APPLICANT: Kosmatka, Anna L.

: APPLICANT: Shilling, Lisa K.

: APPLICANT: Stodola, Robert K.

: APPLICANT: Knowles, David J. C.

: APPLICANT: Black, Michael T.

: APPLICANT: Hodgson, John E.

: APPLICANT: Nicholas, Richard O.

: TITLE OF INVENTION: r1ba

: NUMBER OF SEQUENCES: 7

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: Dechert, Price & Rhoads

: STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre

: CITY: Philadelphia

: STATE: PA

: COUNTRY: US

: ZIP: 19103-2793

: COMPUTER READABLE FORM:

: MEDIUM TYPE: Diskette

: COMPUTER: IBM Compatible

: OPERATING SYSTEM: DOS

: SOFTWARE: FastSeq for Windows Version 2.0

: CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/978,458
FILING DATE: 08/9/11,503
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/911,503
FILING DATE: 15-AUG-1997
APPLICATION NUMBER: PCT/US97/14436
FILING DATE: 15-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: Dickinson, Q. Todd
REGISTRATION NUMBER: 28,354
REFERENCE/DOCKET NUMBER: P50533-09
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215/994-2252
TELEFAX: 215/994-2222
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3417 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-978-456-7

Query Match 10.9% Score 112.6; DB 4; Length 3417;

Best Local Similarity 49.2% Pred. No. 3,4e-19;
Matches 386; Conservative 0; Mismatches 389; Indels 10; Gaps 3;

QY 1 ATGATTTATGCGATTTCACCTTGCATAATATGTTACAGGTCAACAGGTGTTATCCACCC 60
DB 19 ATGAATATGCAATTAATAAGGCGGCTGTTACGTCATCCCAATCCATATG 78
QY 61 GTTGCGCGCTGTGTAGTATGAAGTACGATTTGTGTATGTCGACACTTACGAAAA 120
DB 79 GTTGCGCGCAATTTATGTAAGATATCACTTATCGACAGGTATCATGAGTTTTT 138
QY 121 GGTGACACAGCATGGCGGTTCAAGCATGATATGTCACAAACAAATGCTGAAGTGGC 180
DB 139 GGTGCGCGCATGCTGACAGAAATGCTTTAAAACTGTGAAAAATCCCTGTGGAGCG 198
QY 181 ACGATTTATATACGTAGAGCCATGATGATTTGTTTCACACGACCCCTGTGTAC 240
DB 199 ACGTTTATGTAACACTTGAACCTGTTGTCACCTTCGGAAGAACACCTCCCTGTATGAT 258
QY 241 AAAATTTATGTTGTAAAGTACGAAAAGTATACGCAACAAAGA-----CAATT 292
DB 259 GCTATATCGATAGTGTATATACAGAGTAGTCAATTGGAAGCCTGATGATCTATT 318
QY 293 CGTTAGACACACATGATGAGAGGTACGGGCTCAAGGATTTGAGGTTGATGCGTTG 352
DB 319 GTATCTGAAAAAGAGTAAGATCTTGAGAAAAATATCTTCAAGTTACTGTTGAATT 378
QY 353 ATGATGAACGGGATCAATATATACCAAGACTTTTAAAGCA-AAAGCAAGCAACTG 411
DB 379 TTGAAATAGTGTCTTACTATAATAAAGTTTGAAGATATATACCCACAGTA 438
QY 412 CCACAAATACAGTGAAGTATCTGCAAGTTTACATGTAACAGGAAATGATATGGA 471
DB 439 CCGTATGTTTTTATGAATATGCAATGTCAATGATGTAAGCAACCAAT 498
QY 472 CAAAGTCAATGATTAATCAACAAAGAGTTAAACAAGATGCTTAAGTAAAGCATG 531
DB 499 CAATCCAAATGATTAATGAAGAAAGCAAGAAAGATGCAATGCAAGTACGACAT 558
QY 532 CACGACGAGTGTAACTGGAAGAGTACAGTTGAATTAAGTATGATCAATATAC 591
DB 559 GTTAGTCATTAATGAGGAGTCAATATCTTATCAAGACGATCCGTTGTCATGAT 618
QY 592 CGTATTCAGATGAAAAACCTATATAAAGTAAATATGTTAGTGGGAATATTCAT 651
DB 619 AGATTGAGAGGAAAAATCCTATCCGTTATCATATGCGATACACATTTACGAATCCT 678

QY 652 TTAA-ATCAGCAAAATTTATCAAGATGAATCAACCAATTTGGATATATATCAAAAATCC 710
DB 679 CTTAACCTTAATAATCGTAAAAACAGCAAAATGATATTAATAACTTATGCACTTCTCT 738
QY 711 AAATTTACAGCAATCAACATATGTAATTTACTTGAAGTCTTGTGATTTAC 770
DB 739 GAAGCAAAAATTAATGAAGCTATATCAAAAATCAATGAGTGTGAATATCTTCCATTAAG 798
QY 771 AACA 775
DB 799 AAAAA 803

RESULT 6

US-08-978-454-7
Sequence 7, Application US/08978454
Patent No. 6017728

GENERAL INFORMATION:
APPLICANT: Palmer, Leslie M.
APPLICANT: Fedon, Jason C.
APPLICANT: Warren, Richard L.
APPLICANT: Kosmatka, Anna L.
APPLICANT: Shilling, Lisa K.
APPLICANT: Stodola, Robert K.
APPLICANT: Knowles, David J. C.
APPLICANT: Black, Michael T.
APPLICANT: Hodgson, John E.
APPLICANT: Nicholas, Richard O.
TITLE OF INVENTION: ribh
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dechert, Price & Rhoads
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103-2793
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,454

FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/911,503
FILING DATE: 15-AUG-1997
APPLICATION NUMBER: PCT/US97/14436
FILING DATE: 15-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: Dickinson, Q. Todd
REGISTRATION NUMBER: 28,354
REFERENCE/DOCKET NUMBER: P50533-11
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215/994-2252
TELEFAX: 215/994-2222
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3417 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-978-454-7

Query Match 10.9% Score 112.6; DB 5; Length 3417;
Best Local Similarity 49.2% Pred. No. 3,4e-19;
Matches 386; Conservative 0; Mismatches 389; Indels 10; Gaps 3;

QY 1 ATGATTTATGCGATTTCACCTTGCATAATATGTTACAGGTCAACAGGTGTTATCCACCC 60

```

Db 19 ATGAATATAGCAATAAACTGCACAAAAGGGGCTGTACGTCAATCCCAATCCTATG 78
Qy 61 GTTGGCGCTGTGTATGATGAGGTAGGATGTGTGATGTGTCACACTTGAAGAAA 120
Db 79 GTTGGCGCAATTTATGTAAGATTAATCACTTATGCAAGGTTATATATGATTTTTT 138
Qy 121 GGTGACAAACATGCGGAGGTTCAAGCACTTGATGTCGCAACAAATGCTGAGGTGG 180
Db 139 GGTGGCCCATGCTGAGAGAAATGCTTTAAACTGTAGAAAATCCCTGTGCGAGGG 198
Qy 181 AGCAATTTATTTACGTTAGAGCCAGTGTATCATTTTGTTCACACACCCTGTGTTAAC 240
Db 199 AGCCTTTATGTAACTTAACCTTGTCTCTCACTTGGGAAAAACACCTCCGTATAGAT 258
Qy 241 AAAATATATGATTTAGATAGCAAAAGTAGTATAGCAACAAAAG-----CAAT 292
Db 259 GCTATATATGATGATGATTTACAAAGATGATGATGAAAGCTTACCTGATCTTAT 318
Qy 293 CGTTAGACACATGATGATGAGACGTTACGGGCTCCGATATGAGTTGATGCGTTG 352
Db 319 GTATCTGGAAGGAGATTAAGATTAAGTGAAGAAATATCTTCACTTACTGTGGAAT 378
Qy 353 ATGATGACGGGCATCACAATTTATACCAAGACTTTTAAAGCA-AAAGCAAGCACTG 411
Db 379 TTAGAAAATGATGCTTAACTTAATAAAAGTTTGAAGAGTATTTACCCAGCATGTA 438
Qy 412 CCACAATATACGATGAATATCTGCAAGTTAGATGTAACCAAGCAATGATATGA 471
Db 439 CCTATGATTTTATGAAATATGCAATGCAATGATGAGAAATATACCAATTAACAAAT 498
Qy 472 CAAAGTCATGATTTACTTAACAAGAGTTTAAACAAGATGCTATAGTTAAGACATCGA 531
Db 499 CAATCCAAATGATTTCTGAGAGAGAGCAAGAACATGTCATGATGATAGCACTAT 558
Qy 532 CACGACGACGTGTTAACTGGAAGACGTACAGTTGAATTTGATGATCCACAATATCTACA 591
Db 559 GTTAGTCATTAATGCTGGAGAGTCATTAAGTATTCAGACAGATCTTTGCTGACATGT 618
Qy 592 CATTATCAAGATGGAAGAAACCTATTAAGATATTTGTCTAAGTCTGGATATATCAT 651
Db 619 ACATTTGAGAGAGGAGAAATCCTATCCGTATCATATGCGATACACATTTAGCACTCT 678
Qy 652 TTTA-ATCAGCAATTTATCAAGATGAATCAACACCAATTTGATTTATATGAAATCC 710
Db 679 CTAACTCTTAAATCTGTAAGAACGCAATGATTTAAACCTTACATTTGCTGCTCT 738
Qy 711 AAATTTACAGCAATCAACACATTTGAATTTATTTACTGAGTCTTGTGATTTAAC 770
Db 739 GAAGACAAAATTAATGAAGCTATATCAAAATCATGCTGTGAATACTTCCATTAAG 798
Qy 771 AACAA 775
Db 799 AAAAA 803

```

RESULT 7
 US-08-741-327E-14
 ; Sequence 14, Application US/08/41327E
 ; Patent No. 5925354
 ; GENERAL INFORMATION:

; APPLICANT: Fuller, et al.
 ; TITLE OF INVENTION: Rhinoflavin mutants as vaccines
 ; TITLE OF INVENTION: against Actinobacillus pleuropneumoniae
 ; NUMBER OF SEQUENCES: 17
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: G. Kenneth Smith
 ; STREET: 300 S. Wacker Drive
 ; CITY: Chicago
 ; STATE: IL
 ; COUNTRY: USA
 ; ZIP: 60606
 ; COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/741,327E
; FILING DATE: October 28, 1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: G. Kenneth Smith
; REGISTRATION NUMBER: 43,135
; REFERENCE/DOCKET INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-9808
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1230 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; US-08-741-327E-14

```

Query Match 7.4%; Score 76.2; DB 4; Length 1230;
 Best Local Similarity 47.6%; Pred. No. 2.6e-10;
 Matches 293; Conservative 0; Mismatches 313; Indels 9; Gaps 2;

```

Qy 9 TGGCATTCACCTTGCAATATGTCATCAAGTCACAAACAGCTGTATATCCACCGTTGGCC 68
Db 165 TGCATTCACCTGCGCAAAACAGGTTTAGCTGACGATCCCAATCCGCTTGTGCTG 224
Qy 69 TGTATGATTAATGAAGTAGGATTTGTTATGTCACACTTGAGAAAAGGTGACAA 128
Db 225 TGTATTTGTCAAAACGTTGAATGCTTGGCGAAGTTACATGAAAAGATTTGGTGATG 284
Qy 129 GCATCGGAGGTTCAAGCATTTGATATGCAACAACAAATGCTGAGTGGCAGATTTA 188
Db 285 GCATCGGAGAGTAATGCGCTTTTACATTTAGAGAAATCTTTCCGGGCGACTGCTTA 344
Qy 189 TATTACGTTAGCCATGATGATTTGTTCAACACACCGCTGTGTAACAAATAT 248
Db 345 TGTAAAGCTTGAGCTTTGTCATCACGCGCGACGCGCTTGTGGAATTTATTAAT 404
Qy 249 TGATTTGAAGTAGCAAAAGTAGTATACGCAACAAAGA---CAATTCGTTAGACACACA 305
Db 405 TGAACGAGCATTAATAAAGTATTTATGCTTCAGCGATCCGAATCCTTAGTAGCAGG 464
Qy 306 TGGTG-----ATGAGACGTTACGGGCTCACGATTTGAGGTTGATGCGTTGATGATGA 359
Db 465 GCGGGAGCAAAATCAGCTACGCCCAAGCCGGCGTGGAAGTGGGAGGTTTACTCAAGA 524
Qy 360 ACGGCATCACATTTATACCAAGACTTTTAAAGCAAAACGAAGCAAGTCCACAAAT 419
Db 525 AGAATGTATGCTTAAACCGGATTTTTCACATATTTCAACTAATCAAGCTCCGTATGT 584
Qy 420 TACAGTGAAGATCTGCAAGTTAGATGTAACAACGAATGATGAGCAAAAGTCA 479
Db 585 GCTATATGAATATGCCATGACGCGCAGAGCAAAATTTGCAACCGGTACGGCGAATCCA 644
Qy 480 ATGATTTACTAACAAGAGGTTAAACAAGATGCTATATAGTTAAGACATGCACAGCAGC 539
Db 645 ATGATTTACGGTGAATGGCAAGAGAGTGAGCAAAACAGTCAATATAGTGC 704
Qy 540 ACTGTTAACTGGAAGACGTACAGTTGATTTGATGATCCACAATATATCTAACGATTTCA 599
Db 705 GATTTATGCTGCTGTAGATAGGACTTGCAGATTAACCCGATGTTAATATACCGAATGCC 764
Qy 600 AGATGAAAAAACCC 614
Db 765 GAATCGAACAACCC 779

```


RESULT 8
US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-Fls
US-08-232-463-14

Query Match 5.4%; Score 55.6; DB 1; Length 7218;
Best Local Similarity 7.28; Pred. No. 5.8e-05;
Matches 31; Conservative 220; Mismatches 179; Indels 0; Gaps 0;

229 CCCTGTGTTAAACAAATTTGATGTTAGATGCAAAAGTAGTATACGCAACAAAGAC 288
1487 CACTGTAATTAACCTATCTATGACAGTAGTTAAAGAGATGAGAAATTTGTACRRRRRR 1428
289 AATTCCTGTAGACACACATGATGATGAGCTGACGGCTCAGCGTATGAGTTGAATG 348
1427 RRR 1368
349 GTTGATGATGACGGGCATCATCAATTATACCAAGACTTTTAAAGCAAAAGCAACAA 408
1367 RRR 1308
409 CGGCCCAATTAATGATGAGATGATGCAAGTTAGATGTTAAGAGCAATGATAT 468
1307 RRR 1248
469 GGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 528
1247 RRR 1188

529 CGACAGCAGCAGTGTATGAGAGAGTACGTTGATTAATGATGATGATGATGATGATGAT 588
1187 RRR 1128
589 ACACGTATTCAGATGAGAAAAACCTATAAAGTAAATTTGCTTAAGTCTGGCAATAT 648
1127 RRR 1068
649 CATTTTAATC 658
1067 RATCGCAAGC 1058

RESULT 9
5231168-1
Patent No. 5231168
APPLICANT: DIEGIEL, MORTEN; BORRE, MARTIN; JEPSEN, SOREN;
VUUST, JENS; RIENECK, KLAUS; WIND, ANNETTE; JAKOBSEN, PALLE H.
TITLE OF INVENTION: MALARIA ANTIGEN
NUMBER OF SEQUENCES: 19
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/409,658
FILING DATE: 18-SEP-1989
SEQ ID NO: 1
LENGTH: 3095
5231168-1

Query Match 5.0%; Score 51.2; DB 7; Length 3095;
Best Local Similarity 41.6%; Pred. No. 0.00055;
Matches 323; Conservative 0; Mismatches 453; Indels 0; Gaps 0;

250 GATTGATGATGCAAAAGTATATACGCAACAAAGCAATTCCTTTAGACACATGCT 309
1408 gaagataaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaat 1467
310 GATGAGACGTTGCGGCTCAGCGTATGAGTTGATGAGTTGATGATGATGATGATGAT 369
1468 gataaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaat 1527
370 CAATTATACCAAGACTTTTAAAGCAAAAGCAACCACTGCCAATTAACAGTGA 429
1528 aaaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaat 1587
430 GTATCTGCAAGTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 489
1588 aatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaat 1647
490 AACAAAGAGCTTAACAGATGCTATATAGATGATGATGATGATGATGATGATGATGAT 549
1648 aatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaat 1707
550 GGAAGACGTACGTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 609
1708 gaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaat 1767
610 AACCTATTAAGTATATATGCTTAAGTCTGGAATATTCATTTTAAACGCAATTTAT 669
1768 aagatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaat 1827
670 CAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 729
1828 gaagatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaat 1887
730 ACACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 789
1888 aagatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaat 1947
790 TATTAAGAGGAGTTGGAATTTGCTAGTCCAGGACGATGATGATGATGATGATGATGAT 849
1948 aatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaat 2007

Oy	460	AATGATATGACCAAAATCATGATGATCTACTACAAAGAGSTTAAACAAAGATGTCATTAAG	519
Db	6941	AGTATTTTGAAGTATAGTGAACATTAATTAATTTATTTTCATTAAGATAAAAAATTTAAT	7000
Oy	520	TTAAGACATGCACCGACGATGTTAACTGGAAGCGTACAGTGAATTAAATGATCCA	579
Db	7001	TTAAACCTTAAGATGTTATTGAATTAATGATGAATAATTTAGAAATTTACAAAAATTA	7060

	Query Match	4.5%;	Score 46.4;	DB 1;	Length 5589;
	Best Local Similarity	47.1%;	Pred. No. 0.01;	301;	Mismatches 9;
	Matches 276;	Conservative	0;	Indels	4;
OY	31 GTACAGCTCAAAACGCGTTAATCCACCCTGTGGCCTGTGTGACTT---	AATCAGAGT	87		

Db 672 GTGCGCGGAGTACACAGCCGGAATCCACCGGTGGTGGTAAATTATTCGACTTCGGGT 731
QY 88 AGGATGTGTGTAATGTGTGACACTTGAGAAAAGGTGACAAAGATGGCGAGTTACAGCA 147
Db 732 GAGATTGTGGGACCGGTGACTCAGCCGGTGGGGGTGGCGAGAGAGTCAAGCT 791
QY 148 CTGATATGCGACAAACAAATGCTGAGAGTGCAGCATTTATATATACGTTAGAGCCATGT 207
Db 792 CTAGCGCATCCCGGGGACAGACCGAAGGCGCTACCCCGTGTGAGACGCTGGAGCCCTGC 851
QY 208 AGTCATTTGTTTCACACACACCCCTGTGTACAAATTTATGATTGTAAAGTAGCAAAA 267
Db 852 CGGCAATACCGCGGACGCGGAGCGTGCACCGAGCTTTAATTGAAGCGCGGATCAAGAT 911
QY 268 GTAGTATACCAACAAAGAC---AATTGTTAGACACATGTTGATGACGTTACGG 324
Db 912 GTCTTTTCTTACACTCCGACCCGAAATCCAGTGGCGGGGGAACAGGTGCTTGTGC 971
QY 325 GCTCAGCGTATTGAGT--TGAATGCGTGTGATGATGAGAGGCGATCACAATTATACCAAG 382
Db 972 GATGCGCGCATCAACGTCGTGACGCTGCCAGCCCGGAGGGGTACCGGATCGGCTCAT 1031
QY 383 ACTTTTTTAAAGCAAAAGCAA-AGCAACTGCCACAAATTAACAGTGAAGTATCTGCAAGT 441
Db 1032 CCGTGTGTTGAAGTCCGTCGAGTTCGACGTCGCCACGTTACGTTGAATTTGCGCAACT 1091
QY 442 TTGATGCTAAACAGGATGATTAAGACAAAGTCAATGATTTACTAACAAAGAGTT 501
Db 1092 ATGATGCTTTACTGAGCTGCGATGCGAGTGCACCAAGGAGTACTGGGAAATGGCG 1151
QY 502 AAACAGATGCTATTAAGTAAAGACATGACAGACGACGACGATGTTAATGGAAGACATCA 561
Db 1152 CGGACATGTCACCGGTGACCGTGAACACCGGATGCCATTTATCATCGGACATCGGACG 1211
QY 562 GTTGATTTAGATGATCCACATATACTACACGATTTCAAGATGCA 607
Db 1212 GCCTTGTATGATATCATCGCTTACTGCCGCTATCCAGACGCA 1257

RESULT 12
US-08-487-826B-13
; Sequence 13, Application US/08487826B
; Patent No. 5993827
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,826B
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelien, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CPI

TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19124 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-487-826B-13

Query Match 4.2%; Score 43.4; DB 4; Length 19124;
Best Local Similarity 46.0%; Pred. No. 0.082;
Matches 188; Conservative 0; Mismatches 216; Indels 5; Gaps 1;

QY 390 TAAAGCAAAAGCAACCACTGCCACAATTTACAGTGAAGTATCTGCAAGTTACATGG 449
Db 15493 TAAAAAAG 15552
QY 450 TAAACAGCGAATGATGATGACCAAGTCAATGATTTACTAACCAAGAGGTTAAACAGA 509
Db 15553 AATATATAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 15612
QY 510 TGTCTTAAGTATGACATGACGACGACGACGAGTGTAACTGGAAGCGTACGATTAAT 569
Db 15613 ATTTTAT 15667
QY 570 AGATGATCCCAAT 629
Db 15668 AAAAAATTTAAAAAAT 15727
QY 630 GTCTAAGCTGTGGAATATTTCAATTTAATGACCAATTTAATCAAGATGAATCAACACCA 689
Db 15728 TTAATAAT 15787
QY 690 TTGATATATATGAGAAATCCAAATTTAACAAGCAATCAACACATATGAAATTTTAA 749
Db 15788 AT 15847
QY 750 CTGAGCTGTTGATTTAACAACATTTCTTCAATTTATATATATATATATATATAT 798
Db 15848 ATTAATAAT 15896

RESULT 13
PCT-US93-07261-10
; Sequence 10, Application PC/TUS9307261
; GENERAL INFORMATION:
; TITLE OF INVENTION: P1EMP3 MALARIA ANTIGEN, ANALOGS, ANTIBODIES AND USES THEREO
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John H. C. Blasdale
; STREET: One Giralda Farms
; CITY: Madison
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07940-1000
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 6.0.5
; SOFTWARE: Microsoft Word 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07261
; FILING DATE: 19930805
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/927,531
; FILING DATE: 07-AUG-1992
; ATTORNEY/AGENT INFORMATION:

NAME: Blasdale, John H. C.
REGISTRATION NUMBER: 31,895
REFERENCE/DOCKET NUMBER: DX0288K
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-822-7398
TELEFAX: 201-822-7039
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 4766 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORGANISM: Plasmodium falciparum
STRAIN: Malayan Camp
IMMEDIATE SOURCE:
CLONE: p2bl:p12-1
FEATURE:
NAME/KEY: CDS
LOCATION: 3..4766
PCT-US93-07261-10

Query Match 4.0%; Score 41.6; DB 6; Length 4766;
Best Local Similarity 46.5%; Pred. No. 0.15;
Matches 134; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

OY 442 TTAGATGTAACAAGCGCATGATATGACAAAGTCATTAATCTACCAAGAGTT 501
DB 189 TTAAGGAAAAAGAAATTAAGAACATTTAAAAAGAGTTGAAATGAGAGAA 248
OY 502 AAACAGATGCTATAGTACGACGACGAGCGAGTCTTAAGTGAAGAGCTACA 561
DB 249 AAAGAGAAAAAATTAATTAAGATATATGATGAAGCATTTAAGATTAAGAAAT 308
OY 562 GTTGAATTAGATGATCCACATATATCTACAGCTATTCAGATGAAAAACCTATAAA 621
DB 309 GATTAATGATGATTAAGAAATAGTCCCAAAAAACCAAGATCCGTGAAAGAGATTGAAA 368
OY 622 GTAATATGTCTAAGTCTGGGAATATCTTTTATTCAGCAATTTTCAAGATCAATCA 681
DB 369 GAAATGGAATGAAAGAGAAATTTTAAAAACAACATTTAAAGATTTATGAGAGAGA 428
OY 682 ACACCAATTTGATATATCTGAAATCCAAATTTAAGACGATCAA 729
DB 429 AAGGAAAAAGAAATTTGATCTACTAGACACTTTAAGAGAGATAAA 476

RESULT 14
US-08-883-795A-36/C
Sequence 36, Application US/08883795A
Patent No. 5985607
GENERAL INFORMATION:
APPLICANT: Delcive, Genevieve
APPLICANT: Awang, Gregor
TITLE OF INVENTION: Recombinant DNA Molecules and Expression
TITLE OF INVENTION: Vectors for Tissue Plasminogen Activator
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: BERESKIN & PARR
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5H 3Y2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/883,795A

FILING DATE: 27-JUN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gravelle, Micheline
REGISTRATION NUMBER: 40,261
REFERENCE/DOCKET NUMBER: 7841-062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 665 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
CLONE: Rh 32
US-08-883-795A-36

Query Match 3.9%; Score 40.4; DB 4; Length 665;
Best Local Similarity 44.6%; Pred. No. 0.16;
Matches 207; Conservative 0; Mismatches 251; Indels 6; Gaps 1;

OY 494 AAGAGCTTAACAGAGTCTCTAATAGTAAAGACATCGACAGCAGCGAGTGAACGTGAA 553
DB 516 AAATATGTAATTTAATTAATCTTAATTAATAATTAATTAATTAATTAATTAATAA 457
OY 554 GACGTACAGTTGAATTAATGATGATCCACAAATTAATACAGATTAATCAAGATGAAAAACC 613
DB 456 TATGTAATTAATAAATATGATATTAATTAACATTTTAATTAATAAATGTAATTAATTAAC 397
OY 614 CTATTAAGTATATGCTTAAGTCTGGGAATATTCATTTTAATCAGCAATTTATCAAG 673
DB 396 ATTTTAATTAATAAATATGATATTAATTAACATTTTAATTAATAAATGTAATTAATTAAC 337
OY 674 ATGATACACACCAATTTGGATATATCTGAAATCCAAATTTAACAAGCAATCAACAC 733
DB 336 ATTTTAATTAATAAATATGATATTAATTAACATTTTAATTAATAAATTTTAATTAATTAAC 277
OY 734 ATAT-----TGAATTAATTAATCTTGAAGTCTGATTTAACAACATTTCTTCAATTT 787
DB 276 ATTTTAATTAATAAATATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 217
OY 788 TATATTAAGAGAGTGGAACTTGTCTAGTCGAGCAGGTCACACGACTTCAAGAT 847
DB 216 ATTTTAATTAATAAATATTTAATTAATAAATTTTAATTAATAAATTTTAATTAATTAAT 157
OY 848 TCTGCATCTATTAATAGATGAATTTCTCTATTTATGCCCCGAATTTAATTTGCGCGAT 907
DB 156 ATTTTAATTAATAAATATTTAATTAATTAATTAATTAATTAATAAATTTTAATTAATTA 97
OY 908 CTGGAATTTATCAATTTTATCAACAAGATGATGATGATGATA 951
DB 96 ATTTTAATTAATAAATATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 53

RESULT 15
US-08-918-914-2/C
Sequence 2, Application US/08918914
Patent No. 5876963
GENERAL INFORMATION:
APPLICANT: Mitchell, Peter
APPLICANT: Hutchinson, Nancy
APPLICANT: Lawton, Michael
APPLICANT: Magna, Holly
APPLICANT: Yocum, Sue
APPLICANT: Murry, Lynn E.
TITLE OF INVENTION: HUMAN NUCLEOTIDE PYROPHOSPHORYLASE
NUMBER OF SEQUENCES: 4

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/918,914
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0369
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 4430 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: 777
CLONE: 422069
US-08-918-914-2

```

```

Query Match          3.8%; Score 39.2; DB 3; Length 4430;
Best Local Similarity 49.5%; Pred. No. 0.57;
Matches 101; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 670 CAAGATGATCAACACCAATTTGGATATATACGAAATCCAAATTTTAAACAAGCAATCAA 729
DB 4242 CTATATTATTAAGCTTAACCTTGATGACACTAGAAATTCGAAGAAATTAAGCAAGAAATTA 4183
QY 730 ACACATATTGAATTAATTTACTTGAAGCTCTTGATTTAACAACAATTCCTCACAATTTA 789
DB 4182 ATTAAGTTATTTCATTTCAAGTAAGTAATGTGTTTATTCTTTATTTCACCAACCATTT 4123
QY 790 TATATAAGAGAGAGTGGAACTTGCTAGTCGAGGCAAGTCCAACTACTTCAGAAATTC 849
DB 4122 TATATGTATGAACTAGCAAGGCTTGCTATGTCTGTGAGCAAGGCAAGAAATTAAT 4063
QY 850 TCCATCTATTATATAGTAATTT 873
DB 4062 GCTTTATATTGCTTATTGATCTT 4039

```

Search completed: November 19, 2000, 05:00:11
Job time: 8848 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 19, 2000, 00:12:31 ; Search time 1232.02 Seconds
(without alignments)
5163.965 Million cell updates/sec

Title: US-08-978-456-1
Perfect score: 1029
Sequence: 1 ATGATATATGCGATTCAACT.....TAACCTTAGCAAGAGATGA 1029

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7189864 seqs, 3091403243 residues

Total number of hits satisfying chosen parameters: 14379728

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_est3:*
- 4: gb_est4:*
- 5: gb_est5:*
- 6: gb_est6:*
- 7: gb_est7:*
- 8: gb_est8:*
- 9: gb_est9:*
- 10: gb_est10:*
- 11: gb_est11:*
- 12: gb_est12:*
- 13: gb_est13:*
- 14: gb_est14:*
- 15: gb_est15:*
- 16: gb_est16:*
- 17: gb_est17:*
- 18: gb_est18:*
- 19: gb_est19:*
- 20: gb_est20:*
- 21: gb_est21:*
- 22: gb_est22:*
- 23: gb_est23:*
- 24: gb_est24:*
- 25: gb_est25:*
- 26: gb_est26:*
- 27: gb_est27:*
- 28: gb_est28:*
- 29: gb_est29:*
- 30: gb_est30:*
- 31: gb_est31:*
- 32: gb_est32:*
- 33: gb_est33:*
- 34: gb_est34:*
- 35: gb_est35:*
- 36: gb_est36:*
- 37: gb_est37:*
- 38: gb_est38:*
- 39: gb_est39:*
- 40: gb_est40:*
- 41: em_estba:*
- 42: em_estfun:*
- 43: em_esthum1:*

- 44: em_esthum2:*
- 45: em_esthum3:*
- 46: em_esthum4:*
- 47: em_esthum5:*
- 48: em_esthum6:*
- 49: em_esthum7:*
- 50: em_esthum8:*
- 51: em_esthum9:*
- 52: em_esthum10:*
- 53: em_esthum11:*
- 54: em_esthum12:*
- 55: em_esthum13:*
- 56: em_esthum14:*
- 57: em_esthum15:*
- 58: em_esthum16:*
- 59: em_esthum17:*
- 60: em_esthum18:*
- 61: em_esthum19:*
- 62: em_esthum20:*
- 63: em_estin1:*
- 64: em_estin2:*
- 65: em_estin3:*
- 66: em_estin4:*
- 67: em_estom:*
- 68: em_estov1:*
- 69: em_estov2:*
- 70: em_estp1:*
- 71: em_estp12:*
- 72: em_estp13:*
- 73: em_estp14:*
- 74: em_estp15:*
- 75: em_estp1:*
- 76: em_estro2:*
- 77: em_estro3:*
- 78: em_estro4:*
- 79: em_estro5:*
- 80: em_estro6:*
- 81: em_estro7:*
- 82: em_estro8:*
- 83: em_estro9:*
- 84: em_estro10:*
- 85: em_estro11:*
- 86: em_estro12:*
- 87: em_estro13:*
- 88: gb_gss1:*
- 89: gb_gss2:*
- 90: gb_gss3:*
- 91: gb_gss4:*
- 92: em_gss1:*
- 93: em_gss2:*
- 94: em_gss3:*
- 95: em_gss4:*
- 96: gb_gss5:*
- 97: gb_gss6:*
- 98: gb_gss7:*
- 99: gb_gss8:*
- 100: gb_gss9:*
- 101: em_gss5:*
- 102: em_gss6:*
- 103: em_gss7:*
- 104: em_gss8:*
- 105: em_gss9:*
- 106: em_gss10:*
- 107: em_gss11:*
- 108: gb_gss10:*
- 109: gb_gss11:*
- 110: em_gss12:*
- 111: gb_gss12:*
- 112: gb_gss13:*
- 113: gb_gss14:*
- 114: gb_gss15:*
- 115: gb_gss16:*
- 116: gb_gss17:*

117: gb-gss18:*
 118: gb-gss19:*
 119: em-gss13:*
 120: gb-gss20:*
 121: gb-gss21:*
 122: gb-gss22:*
 123: gb-gss23:*
 124: gb-gss24:*
 125: em-gss14:*
 126: em-gss15:*
 127: em-gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	70.4	6.8	1094	CNS012F2	AL101513 Drosophila
2	68.2	6.6	420	AV426745	AV426745 AV426745
3	63.2	6.1	1101	CNS0039G	AL063921 Drosophila
4	62.4	6.1	1101	CNS0039G	AL069706 Drosophila
5	61.8	6.0	1001	CNS0155H	AL105023 Drosophila
6	61.8	6.0	1101	CNS00EVL	AL069706 Drosophila
7	61.2	5.9	1101	CNS01807	AL108721 Drosophila
8	60.2	5.9	1101	CNS00LJT	AL068307 Drosophila
9	59.6	5.8	1101	CNS0021J	AL061936 Drosophila
10	58.4	5.7	842	CNS00OAO	AL055007 Drosophila
11	58	5.6	781	CNS009DO	AL053444 Drosophila
12	58	5.6	996	CNS00RFH	AL071063 Drosophila
13	56.8	5.5	1218	BE427889	BE427889 pBR699.1
14	56.6	5.5	1201	CNS015EQ	AL106662 Drosophila
15	56.2	5.5	1101	CNS0039R	AL063932 Drosophila
16	56	5.4	945	CNS04DOK	AL285149 Tetradon
17	56	5.4	959	CNS00655	AL062806 Drosophila
18	55	5.3	592	AQ051567	AQ051567 nbxb0002c
19	55	5.3	639	CNS0180D	AL108375 Drosophila
20	54.8	5.3	337	AQ091504	AQ091504 Rfc02461
21	54.4	5.3	1101	CNS0161V	AL106801 Drosophila
22	53.8	5.2	1101	CNS00LJT	AL078714 Drosophila
23	53.2	5.2	970	CNS0182A	AL108736 Drosophila
24	53	5.2	839	CNS011KX	AL100395 Drosophila
25	53	5.2	1092	CNS002K7	AL175696 Tetradon
26	53	5.2	1225	CNS0161D	AL106171 Drosophila
27	52.8	5.1	1029	CNS012GM	AL174271 Tetradon
28	52.8	5.1	1101	CNS003B4	AL064082 Drosophila
29	52.6	5.1	767	CNS00A0X	AL055924 Drosophila
30	52.4	5.1	576	CNS035N7	AL228940 Tetradon
31	52.4	5.1	1001	CNS0064G	AL062271 Drosophila
32	52.4	5.1	1101	CNS008X3	AL052534 Drosophila
33	52.2	5.1	545	AW447058	AW447058 88154 MAR
34	52	5.1	524	CNS01090	AL167541 Tetradon
35	51.8	5.0	928	CNS00DKY	AL071865 Drosophila
36	51.8	5.0	1101	CNS01219	AL101595 Drosophila
37	51.6	5.0	1101	CNS000B8	AL063522 Drosophila
38	51.4	5.0	1101	CNS0039G	AL063921 Drosophila
39	51.2	5.0	1091	CNS014AC	AL013902 Drosophila
40	51	5.0	612	CNS009AI	AL053330 Drosophila
41	51	5.0	783	CNS011BP	AL100940 Drosophila
42	51	5.0	1101	CNS00E07	AL069440 Drosophila
43	50.8	4.9	1101	CNS00E07	AL069440 Drosophila
44	50.8	4.9	1101	CNS016LI	AL106836 Drosophila
45	50.8	4.9	1183	B11823	B11823 T21J7-Sp6 T

ALIGNMENTS

RESULT 1
 CNS012F2

LOCUS	CNS012F2	1094 bp	DNA	GSS	26-JUL-1999
DEFINITION	Drosophila melanogaster genome survey sequence T7 end of BAC BACN07B02 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL101513	GI:5613124			
VERSION	AL101513.1	GI:5613124			
KEYWORDS	fruit fly.				
SOURCE	Drosophila melanogaster				
ORGANISM	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	Genoscope.				
AUTHORS	Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :				
TITLE	BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr				
JOURNAL	BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr				
COMMENT	- Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billard at CEPH (Centre d'etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBeloBAC11.				
FEATURES	Location/Qualifiers				
SOURCE	1..1094				
	/organism="Drosophila melanogaster"				
	/plasmid="pBeloBAC11"				
	/db_xref="taxon:7227"				
	/clone_lib="DrosBAC"				
	/clone="BACN07B02"				
	/note="end : T7"				
BASE COUNT	466 a	151 c	25 g	256 t	196 others
ORIGIN					
Query Match	6.8%; Score 70.4; DB 121; Length 1094;				
Best Local Similarity	36.1%; Pred. No. 2.4e-06;				
Matches 208; Conservative	90; Mismatches 273; Indels 5; Gaps 2;				
450	TAACAAGCAGATGATATGACAAAGTCATGCTTACTTACAAAGAGGTAAACAGA	509			
524	TAAAAAATTCMMWTTAAACVAMCAWCAWCAWCAWCAWCAWCAWCAWCAWCAW	583			
510	TGCTATAAGTTAAGCATGACAGCGAGCTTACTGGAAGAGCTGACGTGATGAT	569			
584	YWCCTTAATTCGCAACCAAAATTTWCCAAATTTWCCAAATTTWCCAAATTTW	643			
570	AGATGATCCACAATATCTACTACAGTATTCAGATGGAAGAAACCTTATTAAGTATAT	629			
644	TTTTTAAAAAAATTTCCAAAAATTTAATTTATTTAAATTTAATTTAATTTAAT	703			
630	GCTTACAGCTGGGATATATTCATTTTATTCAGCAATTTTATCAAGTGAATCAACCCAT	689			
704	TTTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTT	761			
690	TGGAATATATCTGAAATTTCAATTTTACCAAGCATCAACATATTTGAATATTTA	749			
762	TTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTT	821			
750	CTTGAAGCTCTGTGATTTTACACCAATTTCTTCAATTTTATATAAAGAGAGTGGAC	809			
822	TATTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTT	881			
810	TTTGCTAGTCGAGCAGCTGCAACCACTACTTCCAGATTTCCATCTTTATATGATGA	869			
882	WTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTT	941			
870	ATTTATTTCTATTTATGCGCGGAAATTTATTTGCGGATTTGCAATTTTCAATTTATCA	929			
942	TTTTTAAATTTAAATTTTCTTCAWAAWAAWAAWAAWAAWAAWAAWAAWAAW	1001			

LOCUS	DEFINITION	RESULT
990	ACCAATGATGATGATTGACATACGAGATCCGAAACCAATTTGAAATGTCATCTCCAGTT	989
1002	AAAAAAAAATTTAA--AAWMTAAAAAYCAAAAAMAAATTTCTTTTCCAAA	1058
990	ATTAATCAAAATGTTAAATTAACCTTACGAAAGA	1025
1059	WHAAAAAAHMAVCMWTTTCATCTTTTMAAA	1094
2	AV426745	420 bp mRNA EST 23-MAY-2000
AV426745	Lotus japonicus young plants (two-week old) Lotus japonicus cDNA clone MIM071b02_r_5', mRNA sequence.	
AV426745	AV426745.1 GI:7785990	
EST.		
LOCUS japonicus.		
LOCUS japonicus.		
Eukaryota: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta		
Magnoliophyta: eudicotyledons: core eudicots: Rosidae: eustosids I;		
Fabales: Fabaceae: Papilionoideae: Lotus.		
1 (bases 1 to 420)		
Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.		
Generation of 7137 non-redundant expressed sequence tags from a		
legume, Lotus japonicus		
Legume, Lotus japonicus		
DNB Res. 7 (2), 127-130 (2000)		
20277479		
Contact: Yasukazu Nakamura		
The First Laboratory for Plant Gene Research		
Kazusa DNA Research Institute		
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan		
Email: ynakamu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/		
Location/Qualifiers		
1..420		
/organism="Lotus japonicus"		
/db_xref="taxon:34305"		
/clone="MIM071b02_r"		
/clone_jdb="Lotus japonicus young plants (two-week old)"		
/dev_stage="young plants (two-week old)"		
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:		
3hoI; isolate="Miyakojima MG-20"		
BASE COUNT	98 a 72 c 132 g 118 t	
ORIGIN		
Query Match	6.6%; Score 68.2; DB 18; Length 420;	
Best Local Similarity	55.8%; Pred. No. 6.8e-06;	
Matches 130; Conservative	0; Mismatches 103; Indels	0; Gaps
0Y	40 CAAGAGGCTTAATCCACCGCTTGCGCGCTGTGTAATGAAGAGCTAGATTGTTGT	99
Db	1 CAGACCAGCCCCAATTCATGTTGGGTGTGTAATGTAAGAGATGGAAATGTTGT	60
0Y	100 ATTGCTGCACACTTGAGAAAGGTGACACATGCGAGCTTCAACGACTTGATGCA	159
Db	61 CAAGGCTTCAACCTTAAGACGACGAGCCTCATGCTGAGAGTGTCTCTAAGAGATGG	120
0Y	160 CAACAAATGCTGAAGGTGCGACGATTTATATTACCTTAGAGCCATGTGTCATTTGT	219
Db	121 GGGGATTTGGCGAGAGGCTGCGAGGCTTATGTGAGATTGGAGCCTTGTAATCATTTGT	180
0Y	220 TCAACACACCCCTGTGTAAACAAATTAATGATTGTAAGATGACAAAGTAGT	272
Db	181 AGGACTGCACCTTGTTCGAACTTAATTCGAGCCAAAGTCAAAAAGTGT	233
RESULT	3	
CNS00396/c	1101 bp DNA GSS 03-JUN-1999	
LOCUS		
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC #	
	BACR08K10 of Rpcl-98 library from Drosophila melanogaster (fruit	

ACCESSION	fly), genomic survey sequence.
VERSION	AL063921
KEYWORDS	GSS.
SOURCE	fruit fly.
ORGANISM	Drosophila melanogaster
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
	Pterygota; Neoptera; Endopterygota; Diptera; Brachyceta;
	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE	1 (bases 1 to 1101)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammosser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES	Location/Qualifiers
Source	1..1101
	/organism="Drosophila melanogaster"
	/db_xref="taxon:7227"
	/clone_1lb="RPCI-98"
	/clone="BACR08K10"
	/note="end : TET3"
BASE COUNT	201 a 64 c 131 g 202 t 503 others
ORIGIN	
Query Match	6.1%; Score 63.2; DB 121; Length 1101; Best Local Similarity 17.4%; Pred. No. 0.00011; Matches 110; Conservative 276; Mismatches 244; Indels 4; Gaps 2
OY	189 TATTTCGTTAGAGCATGTAGTCAATTTGGTTCACAACCACCCTGTGTAACAAATTAT 248
Db	1100 KARRRGDDTDWDRTRFKDDMDWTWKMTGWKKDRDRRMAGDADRMAWDGCGTWMTATWWW 1041
OY	249 TGATTGTAAAGATAGCAAAGTAGTATACGCCAACAAAGCAATTGTTTAGACACACATG 308
Db	1040 WWWWMTATDTWDMKKMMWWATAAKDTTATWMTARAWAGAGRDGRAGKRDRDAADADGA 981
OY	309 TGATAGACGTTACCGCGTCACGTAATGAGGTGAAG--CGTTGATGATGAACGGCGCA 366
Db	980 GRBDGRRRKDKDKDDDDKGGKKKKAAPAAWAATKWMDWMDWMDKWKGADKRA 921
OY	367 TCACATTAATACCAAGACTTTTTTAACCAAAAGCAAGCAACTGCCCAATTAACATG 426
Db	920 DDDDSAGSKDDDGKKKADDDTDTGRTKDDDKDKMDWMDKAKGTWGDATWAMAATDMMWVG 861
OY	427 AAAGTATCGCAAGTTTAGATGTTAACAAGCGAATGATATGCAAGAATCAATGATT 486
Db	860 WADADMWTVDAADDMWADDRWDAWAMKWMDAMWAGARTADRDRDWGRGKRGARKRRD 801
OY	487 ACTACCAAGAGGTTAAACAAGATGTCATATAAGTTAAAGACATGCAGACGACGATGTTA 546
Db	800 RKRADKRDADDDRDAATTTTWTTRTDDTMWTKDTWTRMAADRTWRDDEDDDRR 741
OY	547 ACT--GGAAGACGTCAGTGTGATAGATGATCCCAATATCTACACGATATTCAGANG 604
Db	740 AGTAGKKRRRTWKRRWKRRDTRWDDADADDTARDRRRRGDDGADAGCKKTGKKRRRRD 681

```

Query Match          6 1%; Score 62.4; DE 121; Length 1101;
Best Local Similarity 34.5%; Pred. No. 0.0017;
Matches 149; Conservative 78; Mismatches 204; Indels 1; Gaps 1;

Oy 581 AATATACACTACGTATTCACAGATGCAAAAACCCATAAAGTATATGTCTAAGTCG 640
   :::::  | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1099 WWTATATWTTTTTTTTTTTWTATATWTATATATATATATATATTTWTTATATATWTA 1040
   :::::  | | | | | | | | | | | | | | | | | | | | | | | | | | |

Oy 641 GGAATATTCATTATTAATCAGCAAAATTTATCAAGATGATCAACACCAATTTGCAATATA 700
   | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT	5
Locus	CNS0155H
DEFINITION	Drosophila melanogaster genome survey sequence SP6 end of BAC BACN13C23 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION	AL105023
VERSION	AL105023.1 GI:5617037
KEYWORDS	GSS.
SOURCE	fruit fly.
ORGANISM	Drosophila melanogaster
REFERENCE	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 1001)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk . This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelosAC11.
FEATURES	Location/Qualifiers
source	1..1001 /organism="Drosophila melanogaster" /plasmid="pBelosAC11" /db_xref="taxon:7227" /cclone_1lp="DrosBAC" /cclone="BACN13C23" /note="end : SP6"
BASE COUNT	266 a 219 c 134 g 150 t 232 others
ORIGIN	
Query Match	6.0%; Score 61.8; DB 121; Length 1001;
Best Local Similarity	32.7%; Pred. No. 0.00023;
Matches 145; Conservative 84; Mismatches 215; Indels 0; Gaps 0;	
353 ATGATGAACGGGCATCAATTATTCACAAGACTTTTAAAGCAAACCAAGCACACGC 412	

Db 553 AAGRVRARWWDGAGAAGADAGATATNAGAAAAAAGRGMAAABDAGRASAAREKTKN 612

Qy 413 CACAATTACAGGAGAGTCTGCAAGTTAGATGTTAAACAAGCAATGATTAAGAC 472

Db 613 GAGGAGAAAAAGWDADDAATWRRKAAARAKAADWMAAACAANAANVMAAANAAGKA 672

Qy 473 AAGTCATGATGATTAACAAAGAGTTAAACAAGATGCTTAAGTAAAGATGAC 532

Db 673 AARRRAAAAAAARAGDAAAAAAAGRAARDAARAAAAATTAAMWATTAARNAARRDA 732

Qy 533 ACAGACGAGTGTAACTGAGAGACGTACAGTTAGATGATCCCAATATACAC 592

Db 733 RAATAAATTAATWMTAAWMAAAMWMAAATAAATTTATWTAATAAATAAATAA 792

Qy 593 GTATTCAGATGAAAAAACCCTATATAAGTATATGCTAGTCTGGCAATATTCAT 652

Db 793 AAAAATATTTAMWMAAATAAATAAATTTWMAAATAAATAAATAAATAAATTTT 852

Qy 653 TTAATCAGCAATTTATCAAGATGAATCAACACCAATTTGATATACGAAATCCAA 712

Db 853 TAAAMMAAAMWATWMAATTAATAADAAAAAATAAATAAATAAATAAATTTWMT 912

Qy 713 ATTAAACAAGCAATCAACACATATGAAATTTATTTACTTGAAGCTTGTGATTTAA 772

Db 913 AAMADMTTWTWTWMAAATAAATWMTATWTTAATWTAATAAATAAATAAATAA 972

Qy 773 CAATCTCTCAATTTATATAA 796

Db 973 WAAAAWMAATWMTAAAAA 996

RESULT 6

CNS00EVL 1101 bp DNA GSS 04-JUN-1999

LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC: BACN329823 of RPCL-98 library from Drosophila melanogaster (fruit fly); genomic survey sequence.

DEFINITION

ACCESSION AL069706.1 GI:4949849

VERSION

KEYWORDS

SOURCE

ORGANISM

fruit fly.

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

1 (bases 1 to 1101)

AUTHORS

Genoscope.

TITLE

Direct Submission

JOURNAL

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

COMMENT

- web : www.genoscope.cns.fr

- determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see : <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCL-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the library pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source

1..1101

/organism="Drosophila melanogaster"

/db_xref="taxon:7227"

/clone_lib="RPCL-98"

/clone="BACR29823"

/note="end : 77"

BASE COUNT 419 a 91 c 60 g 299 t 232 others

ORIGIN

Query Match 6.0%; Score 61.8; DB 121; Length 1101;

Best Local Similarity 31.8%; Pred. No. 0.0023;

Matches 183; Conservative 95; Mismatches 297; Indels 0; Gaps 0;

Qy 371 AATTATACCAAGACTTTTAAAGCAAAAGCAAGCAACTGCCACAATTTACAGTAAG 430

Db 492 AATWAAAAAAMWMAATTTTWWMTWTTWTTWMAATWMAAATAAATAAATAA 551

Qy 431 TATCTCAGATTTAGATGTAACAAGCAATGATATGACCAAGTCAATGATTA 490

Db 552 TTAAMWMAATTAATTAATAATTTAAAMWMTATATTAATTAATTAATTAATA 611

Qy 491 ACAAGAGGTTAAACAAGATGCTATTAAGTAAAGACGACGACGAGTGTAACTG 550

Db 612 AAAATATTTTWTATTAATTTTAAATTAATTAATTAATTAATTAATTAATTA 671

Qy 551 GAAGACGTACAGTGAATTAAGATGATCCCAATATCTACAGTATCAAGTGA 610

Db 672 ATTAATTAATTTWMAATTAATAAATAAATAAATAAATAAATAAATAAATAA 731

Qy 611 ACCCTAAAGCAATATGCTTAAGTCTGGAATATTTCAATTTATCAAGAAATTTATC 670

Db 732 WAATTAAMMAAAMWATWMAATTAATTAATTAATTAATTTAAWMAATTAATA 791

Qy 671 AAGATCAATCAACACCAATTTGATATTAATCTGAATAATCCAAATTTTAAACA 730

Db 792 WAATTAAMMAAAMWATTAATTAATTAATTAATTAATTAATTAATTAATTAATA 851

Qy 731 CAATATTTGAATTTATTTACTGTAAGTCTTGATTTACACAAATCTTCAATTTAT 790

Db 852 AAAATTAATTAATTAATTAATTAATAAATAAATAAATAAATAAATAAATAA 911

Qy 791 ATAAAGAGAGTTGAGCTTGTCTAGCAGCAGCTCAACCACTACTTCAAGATTTCT 850

Db 912 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 971

Qy 851 CCATCTATATATAGATTAATTTCTATATATGCCCCGAATTAATTTGGCGATCG 910

Db 972 AATWTTAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATAA 1031

Qy 911 GAATATATCAATTTTTCACAACCAATGATGAT 945

Db 1032 TAATATATWATWMTWMAATTAATAAATAAATAAATAAATAAATAAATA 1066

RESULT 7

CNS01807/c

LOCUS Drosophila melanogaster genome survey sequence SP6 end of BAC: BACN3707 of DrosBAC library from Drosophila melanogaster (fruit fly); genomic survey sequence.

DEFINITION

ACCESSION AL108721.1 GI:5629025

VERSION

KEYWORDS

SOURCE

ORGANISM

fruit fly.

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

1 (bases 1 to 1101)

AUTHORS

Genoscope.

TITLE

Direct Submission

JOURNAL

Submitted (23-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

COMMENT

- web : www.genoscope.cns.fr

- determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - <http://www.edgp.ebi.ac.uk> - This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billand at CPH (Centre

d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelosAC11.

FEATURES

Location/Qualifiers
1. 1101

Source
/organism="Drosophila melanogaster"
/plasmid="pBelosAC11"
/db_xref="taxon:7227"
/clone_lib="DrosBAC"
/clone="BACN37F07"
/note="end : Sp6"
BASE COUNT 179 a 70 c 85 g 448 t 319 others
ORIGIN

Query Match 5.9%; Score 61.2; DB 121; Length 1101;
Best Local Similarity 42.5%; Pred. No. 0.00031;
Matches 182; Conservative 21; Mismatches 221; Indels 4; Gaps 1;

Y 371 AATTACCAGACTTTTAAAGCAAAAGCAACGACCTGCCACAATTACAGTGAAG 430
D 562 AAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 503
Y 431 TATCTCAAGTTTACATGTAACAAGCAATGATATGACAAAGTCATGATTACTA 490
D 502 TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 443
Y 491 ACAAGAGGTTAACACAGATGCTCTTAAGTTAGACATCGACACGCGAGTCTTACTG 550
D 442 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 383
Y 551 GAAGCGTACGTTGATTAATAGATGATCCACATATATCTACAGTTCAGAGTGAATA 610
D 362 AATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 323
Y 611 ACCCTATAAAGTATATGCTCTAAGTCTGGAATATTCATTTTATCAGCAAAATTATC 670
D 322 AATAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 267
Y 671 AAGATGAATCAACACCAATTGATATATCTGAAATTCCTCAATTTTAAACAAGCAATCA 730
D 266 AATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 207
Y 731 CACATATGAATTTTACTGTAAGTCTGTGATTTAAACAACATCTCTCCCAATTTAT 790
D 206 WAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 147
Y 791 ATAAAGA 798
D 146 AAAAAAA 139

RESULT 8
CNS001J 1101 bp DNA GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC:
DEFINITION BACR30P10 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL068307
VERSION AL068307.1 GI:4958538
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscophora; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
1 (bases 1 to 1101)
Genoscope.
TITLE Direct Submission
AUTHORS Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutyo Osoegawa and Aaron Mammoeer in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

Source

Location/Qualifiers
1. 1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPCI-98"
/clone="BACR30P10"
/note="end : T7"
BASE COUNT 507 a 148 c 112 g 171 t 163 others
ORIGIN

Query Match 5.9%; Score 60.2; DB 121; Length 1101;
Best Local Similarity 37.0%; Pred. No. 0.00053;
Matches 160; Conservative 54; Mismatches 219; Indels 0; Gaps 0;

Y 369 ACAATTATACCAAGCTTTTAAAGCAAAAGCAACGACCTGCCACAATTACAGTGA 428
D 121 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 180
Y 429 AGTATCGCAAGTTTACATGCTTAACAAGCAATGATTAATGACAAAGTCAATGATTAC 488
D 181 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 240
Y 489 TACAAAGAGTTTAAACAGATGCTTAAGTTAAGCATCGACAGCAGGTGTAAC 548
D 241 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 300
Y 549 TGAAGAGCTGACGTTGATTAATGATGATCCACATATATCTACAGTATTCAGATGAAA 608
D 301 WAAATCAAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 360
Y 609 AAACCTATAAAGTATATTTGCTAAGTCTGGCAATATTCATTTTAAATCAGCAATTTA 668
D 361 WAAWACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 420
Y 669 TCAGATGAATCAACACCAATTGATATATCTGAAATTCACAAATTTAACAAGCAATCA 728
D 421 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 480
Y 729 AACACATATTAATTTTACTGGAAGTCTGTGATTTAACAACAATTCCTCACAATTT 788
D 481 AAAAGAAACMAAAAMWMTAAAGWMAACATMTATACCAAAAAAAGAAACCAACCTTA 540
Y 789 ATATTAAGAGCA 801
D 541 WAAAAAGACRWGA 553

RESULT 9
CNS0021J 1101 bp DNA GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION BACR05N11 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL061936
VERSION AL061936.1 GI:4940214
KEYWORDS GSS.
SOURCE fruit fly.

ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mamoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPc1-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source
1..1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPc1-98"
/clone="BACR05N11"
/note="end : 7c 28 g 289 t 146 others

BASE COUNT 631 a 7 c 28 g 289 t 146 others

ORIGIN

Query Match 5.8%; Score 59.6; DB 121; Length 1101;
Best Local Similarity 38.2%; Pred. No. 0.00074;
Matches 261; Conservative 36; Mismatches 387; Indels 0; Gaps 0;

333 TATGAGTTGATGCGTGTGATGATGACGGGATCACAATTATACCAAGACTTTTAA 392
||||| : : : : :
100 TTTTATTTTAAKAKMPTAATAAANHAANAAGTGTTCATGCAAGTTTAA 159
393 AGCAAAAGCAACCACTGCCACAATTCAGTGAAGTATCTGCAAGTTTACATGTTAA 452
||||| : : : : :
160 AAAAAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 219
453 ACAAGCAAGTGAATGACAAAGTCAATGATTAACAAGAGTTTAAACAATGT 512
||||| : : : : :
220 AAAAAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 279
513 CTATAGTTAGACATGACGACGACGAGTGTAACTGAGAGCACTGATTAAGCA 572
||||| : : : : :
280 AAAAAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 339
573 TGAATCAATATCTACAGTATTCAGATGGAAGCAAGCAAGCAAGCAAGCAAGCAAG 632
||||| : : : : :
340 AAAAAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 399
633 TAAGTGGGAATATCTATTTTATCAGCAATTTTCAAGATGATCAACACCAATTTG 692
||||| : : : : :
400 AAAAAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 459
693 GATATATAGTGAATTCGAATTCGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 752
||||| : : : : :
460 AAAAAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 519
753 GAAGCTTGTGATTTAACAACAATCTTCAACAATTTATATAAGAGAGGAGTTGAACTTT 812
: : : : :
520 TTATATTTTAT 579
813 GCTAGTCGAGGAGTCAACCACTACTCTCATATCTCATATATATATATATATATAT 872

Db 580 AATTTATATAAAMATWTTTATTTATATWTAAMAAAAAAATTTTAAAMATWTTT 639
Qy 873 TATCTCTATATATGCGCAAAATTAATGCGGATCGGAAATATATCAATTTATCAAC 932
Db 640 TTTTATATATATATAAAMATWTTTATATWTAAMAAAAATTTTAA 699
Qy 933 AATGATGATTTGATGATACCGAGATCGCAACATTTGAAATTTGTTCAATCCGAGTAT 992
Db 700 AATTTTAT 759
Qy 993 AATCAAAATGTTAATTAACCTT 1016
Db 760 TAT 783

RESULT 10
CNS00ACO/c
LOCUS
DEFINITION
CNS00ACO 842 bp DNA GSS 04-JUN-1999
Drosophila melanogaster genome survey sequence T7 end of BAC #
BACR21B05 of RPc1-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
AL055007
VERSION
AL055007.1 GI:4936180
KEYWORDS
GSS
SOURCE
fruit fly.
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 842)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mamoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPc1-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source
1..842
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPc1-98"
/clone="BACR21B05"
/note="end : 7c 78 g 44 g 464 t 39 others

BASE COUNT 217 a 78 c 44 g 464 t 39 others

ORIGIN

Query Match 5.7%; Score 58.4; DB 121; Length 842;
Best Local Similarity 45.6%; Pred. No. 0.0014;
Matches 164; Conservative 10; Mismatches 186; Indels 0; Gaps 0;

389 TTAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 448
Db 580 TTTTAAAAAATTAAT 521
Qy 449 GTAACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 508
||||| : : : : :
520 TAAAAAATTAAT 461

QY	509	ATGCTATTAAGTAGACATGCACGCGCGGTGA	CTTAACGTGAAGACGTACAGTTGAAT	568
Db	460	AAAAAAAAATATATAATTAAATTA	AAAAAAAAATATAAATATATAATAATATATA	401
QY	569	TAGATGATCCCAATATCTACGCTATTCAGAT	TGAAAAAAACCCTATAAAGTAAT	628
Db	400	AAATATATTAATTAATAAAAAA	AAAAAAGRAAATAAGAAAAAGRAAARAT	341
QY	629	TGCTTAAGCTTGGCAATATTCATTTT	TATCAGCAATTTATCAAGTGAATCAACCCA	688
Db	340	GRARATAAAGGGRARARAA	CCAAATTTATTTTMAAATTTAAAAATRRAT	281
QY	689	TTTGATATATATCTGAATCCAAATTT	TACAGCAATCAACACATATGGAATTT	748
Db	280	TRATATATMAATGACAAATTAAT	TTTRRAAAATAAAAAATGATTTTAATATATAT	221

RESULT	11
CNS009DO	
LOCUS	761 bp DNA
DEFINITION	Drosophila melanogaster genome survey sequence TERN3 end of BAC # BACR1009 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION	AL053444	GI:4934889
VERSION	AL053444.1	
KEYWORDS	GSS.	
SOURCE	fruit fly.	
ORGANISM	Drosophila melanogaste	

REFERENCE 1 (bases 1 to 781)

TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :

COMMENT

Submitted (02-JUN-1999) Genoscope - Centre National de Séquençage
BP 191 91006 Evry cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org/TheBDGP/Drosophila>
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mamozer in Pierer de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPc1-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y¹; cn bw sp. the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BAC/PC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

Source

BASE COUNT	440 a	189 t	75 others
ORIGIN	71 c	6 g	
	/note="end : TET3"		
	/clone="BACR19009"		
	/clone_lib="RPCT-98"		
	/db_xref="taxon:7227"		
	/organism="Drosophila melanogaster"		

Query Match	5.68;	Score 58;	DB 121;	Length 781;
Best Local Similarity	42.78;	Pred. No. 0.0017;		
Matches 178;	Conservative 18;	Mismatches 221;	Indels 0;	Gaps 0;

Oy 380 AAGACTTTTAAAGCAAAAGCAGCTCCACAAATTACGTAAGAATCTGCAA 4399
||| ||| ||| ||| :||| :||| :||| :|||
Db 117 AAAAATTTAAAAATATAAAAAMWAAAAATATAAMMMAATAMAAAAAWAAAAAA 176
Oy 440 GTTTAGATGTAAACCMCGCATGATTAATGCACAAGTCATGATTACTPACAAGAGG 4999

Db	177	AAAMAAAAAAMAAAAAAMAAAAATATMTAAAAAATTTAAAAATATATATAATAA	236
Qy	500	TTAAACAAGATGCTATTAAGTAAAGCATCGACACGACGAGCTTAACTCGAAGACGTA	559
Db	237	AAAAAAMAAAMWTTTTTAAAAATTAATTTTTTAAAAATTTAAAMAAATTTATTA	296
Qy	560	CAGTTGATTTGATGATCCACAATATCTCACGATATTCAGATGCAAAAACCTATAA	619
Db	297	AAATTTAAAAAAATTAAMAAAAAAMAAAAAAMAAAAAATTAAMAAATTAAMAAAAA	356
Qy	620	AAGTATATTTGCTCTAGCTCGGAGATATTCATTTTAATCAGCAAATTTATCAAGATGAAT	679
Db	357	AAAAAMTAAAAAAMAAAMWTTATTAATAATTTATTAATTAATTAATTAATTA	416
Qy	680	CAACACCAATTTGATATATATCTGAAAAATCCAAATTTAAACAGCAATCAACACATATG	739
Db	417	AAATTAATTAATTAAMAAAAAAMAAAAAATTTAAAAAATTTAAAAAAMAAAAAATTA	476
Qy	740	AAATTAATTAAGCTGTGATTTTACACAAATCTTCACAATTTATTTATTA	796
Db	477	ATAATTAATAATAATAATAACAAAAAATTAATTAATAATAATAATTAATTTAAAAAT	533

RESULT	12
CNS00FUH	
LOCUS	CNS00FUH 996 bp DNA- 03-JUN-1999
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC:

ACCESSION	AL071063
VERSION	AL071063.1
GI	GI:4951105

SOURCE ORGANISM

REFERENCE 1 (bases 1 to 996)

TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :

COMMENT

Submitted (02-JUN-1999) Genoscope - Centre National de Séquençage
BP 191 91006 Evry cedex - FRANCE (E-mail : segrete@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org/TheBDGP/Drosophila>
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mamooser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPc1-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y¹; cn bw sp. the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

Source

BASE COUNT	383 a	164 c	81 g	171 t	197 others
------------	-------	-------	------	-------	------------

Query Match	5.6%	Score 58:	DB 121:	Length 996:
Best Local Similarity	34.0%	Pred. No.	0.0017:	
Matches 177:	Conservative	81:	Mismatches	260:
			Indels	2:
			Gaps	1:

FEATURES	source
ORGANISM	Barley.
ORGANISM	Eukarya; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Hordeum.
REFERENCE	Anderson, O.A., Appels, R., Bailey, P., Blake, T., Close, T., Cloutier, S., Dubcovsky, J., Feuillet, C., Gale, M., Graner, A., Gustafson, P., Herrmann, R.G., Holton, T., Jacquemin, J.M., Jia, J., Joudrier, P., Langridge, P., Lazo, G.R., Lin, J.J., McGuire, P., Ogihara, Y., Peckol, N., Qualset, C., Schuch, W., Selvaraj, S., Sharifou, M., Sorrells, M., Warburton, M., and Wenzel, G.
TITLE	International Triticace EST Cooperative (ITREC): Production of Expressed Sequence Tags for Species of the Triticaceae Unpublished (2000)
JOURNAL	Contact: Bailey P
COMMENT	Cereals Group, John Innes Centre Norfolk, Norwich NR4 7UH UNITED KINGDOM Tel: 44 1603 452571 ext. 2587 Fax: 44 1603 502241 Email: paul.bailey@bbsrc.ac.uk International Triticace EST Cooperative (ITREC) http://wheat.pw.usda.gov/genome. Location/Qualifiers 1. 1218
FEATURES	source
LOCUS	BEA27889 1218 bp mRNA EST 24-JUL-2000
DEFINITION	PSR6899 ITREC PSR Barley Inflorescence Library Hordeum vulgare cDNA clone PSR6899, mRNA sequence.
ACCESSION	BEA27889
VERSION	BEA27889.1 GI:9425732
KEYWORDS	EST.
SOURCE	Barley.
ORGANISM	Eukarya; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Hordeum.
REFERENCE	Anderson, O.A., Appels, R., Bailey, P., Blake, T., Close, T., Cloutier, S., Dubcovsky, J., Feuillet, C., Gale, M., Graner, A., Gustafson, P., Herrmann, R.G., Holton, T., Jacquemin, J.M., Jia, J., Joudrier, P., Langridge, P., Lazo, G.R., Lin, J.J., McGuire, P., Ogihara, Y., Peckol, N., Qualset, C., Schuch, W., Selvaraj, S., Sharifou, M., Sorrells, M., Warburton, M., and Wenzel, G.
TITLE	International Triticace EST Cooperative (ITREC): Production of Expressed Sequence Tags for Species of the Triticaceae Unpublished (2000)
JOURNAL	Contact: Bailey P
COMMENT	Cereals Group, John Innes Centre Norfolk, Norwich NR4 7UH UNITED KINGDOM Tel: 44 1603 452571 ext. 2587 Fax: 44 1603 502241 Email: paul.bailey@bbsrc.ac.uk International Triticace EST Cooperative (ITREC) http://wheat.pw.usda.gov/genome. Location/Qualifiers 1. 1218
FEATURES	source
LOCUS	BEA27889 1218 bp mRNA EST 24-JUL-2000
DEFINITION	PSR6899 ITREC PSR Barley Inflorescence Library Hordeum vulgare cDNA clone PSR6899, mRNA sequence.
ACCESSION	BEA27889
VERSION	BEA27889.1 GI:9425732
KEYWORDS	EST.
SOURCE	Barley.
ORGANISM	Eukarya; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Hordeum.
REFERENCE	Anderson, O.A., Appels, R., Bailey, P., Blake, T., Close, T., Cloutier, S., Dubcovsky, J., Feuillet, C., Gale, M., Graner, A., Gustafson, P., Herrmann, R.G., Holton, T., Jacquemin, J.M., Jia, J., Joudrier, P., Langridge, P., Lazo, G.R., Lin, J.J., McGuire, P., Ogihara, Y., Peckol, N., Qualset, C., Schuch, W., Selvaraj, S., Sharifou, M., Sorrells, M., Warburton, M., and Wenzel, G.
TITLE	International Triticace EST Cooperative (ITREC): Production of Expressed Sequence Tags for Species of the Triticaceae Unpublished (2000)
JOURNAL	Contact: Bailey P
COMMENT	Cereals Group, John Innes Centre Norfolk, Norwich NR4 7UH UNITED KINGDOM Tel: 44 1603 452571 ext. 2587 Fax: 44 1603 502241 Email: paul.bailey@bbsrc.ac.uk International Triticace EST Cooperative (ITREC) http://wheat.pw.usda.gov/genome. Location/Qualifiers 1. 1218
FEATURES	source
LOCUS	BEA27889 1218 bp mRNA EST 24-JUL-2000
DEFINITION	PSR6899 ITREC PSR Barley Inflorescence Library Hordeum vulgare cDNA clone PSR6899, mRNA sequence.
ACCESSION	BEA27889
VERSION	BEA27889.1 GI:9425732
KEYWORDS	EST.
SOURCE	Barley.
ORGANISM	Eukarya; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Hordeum.
REFERENCE	Anderson, O.A., Appels, R., Bailey, P., Blake, T., Close, T., Cloutier, S., Dubcovsky, J., Feuillet, C., Gale, M., Graner, A., Gustafson, P., Herrmann, R.G., Holton, T., Jacquemin, J.M., Jia, J., Joudrier, P., Langridge, P., Lazo, G.R., Lin, J.J., McGuire, P., Ogihara, Y., Peckol, N., Qualset, C., Schuch, W., Selvaraj, S., Sharifou, M., Sorrells, M., Warburton, M., and Wenzel, G.
TITLE	International Triticace EST Cooperative (ITREC): Production of Expressed Sequence Tags for Species of the Triticaceae Unpublished (2000)
JOURNAL	Contact: Bailey P
COMMENT	Cereals Group, John Innes Centre Norfolk, Norwich NR4 7UH UNITED KINGDOM Tel: 44 1603 452571 ext. 2587 Fax: 44 1603 502241 Email: paul.bailey@bbsrc.ac.uk International Triticace EST Cooperative (ITREC) http://wheat.pw.usda.gov/genome. Location/Qualifiers 1. 1218
FEATURES	source
LOCUS	BEA27889 1218 bp mRNA EST 24-JUL-2000
DEFINITION	PSR6899 ITREC PSR Barley Inflorescence Library Hordeum vulgare cDNA clone PSR6899, mRNA sequence.
ACCESSION	BEA27889
VERSION	BEA27889.1 GI:9425732
KEYWORDS	EST.
SOURCE	Barley.
ORGANISM	Eukarya; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Hordeum.
REFERENCE	Anderson, O.A., Appels, R., Bailey, P., Blake, T., Close, T., Cloutier, S., Dubcovsky, J., Feuillet, C., Gale, M., Graner, A., Gustafson, P., Herrmann, R.G., Holton, T., Jacquemin, J.M., Jia, J., Joudrier, P., Langridge, P., Lazo, G.R., Lin, J.J., McGuire, P., Ogihara, Y., Peckol, N., Qualset, C., Schuch, W., Selvaraj, S., Sharifou, M., Sorrells, M., Warburton, M., and Wenzel, G.
TITLE	International Triticace EST Cooperative (ITREC): Production of Expressed Sequence Tags for Species of the Triticaceae Unpublished (2000)
JOURNAL	Contact: Bailey P
COMMENT	Cereals Group, John Innes Centre Norfolk, Norwich NR4 7UH UNITED KINGDOM Tel: 44 1603 452571 ext. 2587 Fax: 44 1603 502241 Email: paul.bailey@bbsrc.ac.uk International Triticace EST Cooperative (ITREC) http://wheat.pw.usda.gov/genome. Location/Qualifiers 1. 1218
FEATURES	source
LOCUS	BEA27889 1218 bp mRNA EST 24-JUL-2000
DEFINITION	PSR6899 ITREC PSR Barley Inflorescence Library Hordeum vulgare cDNA clone PSR6899, mRNA sequence.
ACCESSION	BEA27889
VERSION	BEA27889.1 GI:9425732
KEYWORDS	EST.
SOURCE	Barley.
ORGANISM	Eukarya; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Hordeum.
REFERENCE	Anderson, O.A., Appels, R., Bailey, P., Blake, T., Close, T., Cloutier, S., Dubcovsky, J., Feuillet, C., Gale, M., Graner, A., Gustafson, P., Herrmann, R.G., Holton, T., Jacquemin, J.M., Jia, J., Joudrier, P., Langridge, P., Lazo, G.R., Lin, J.J., McGuire, P., Ogihara, Y., Peckol, N., Qualset, C., Schuch, W., Selvaraj, S., Sharifou, M., Sorrells, M., Warburton, M., and Wenzel, G.
TITLE	International Triticace EST Cooperative (ITREC): Production of Expressed Sequence Tags for Species of the Triticaceae Unpublished (2000)
JOURNAL	Contact: Bailey P
COMMENT	Cereals Group, John Innes Centre Norfolk, Norwich NR4 7UH UNITED KINGDOM Tel: 44 1603 452571 ext. 2587 Fax: 44 1603 502241 Email: paul.bailey@bbsrc.ac.uk International Triticace EST Cooperative (ITREC) http://wheat.pw.usda.gov/genome. Location/Qualifiers 1. 1218
FEATURES	source
LOCUS	BEA27889 1218 bp mRNA EST 24-JUL-2000
DEFINITION	PSR6899 ITREC PSR Barley Inflorescence Library Hordeum vulgare cDNA clone PSR6899, mRNA sequence.

```

/clone="PSR6899"
/clone_lib="JREC_PSR_Barley_Inflorescence_Library"
/rlsuse_type="Inflorescence"
/dev_stage="Immature_developing"
/note="Vector: Lambda ZAP II (Stratagene)"

BASE COUNT      201 a      160 c      158 g      530 t      169 others

ORIGIN

Query Match      5.5%; Score 56.8; DB 35; Length 1218;
Best Local Similarity 38.3%; Pred. No. 0.0033;
Matches 174; Conservative 0; Mismatches 279; Indels 1; Gaps 1;

OY      356      ATGAACGGGCATCACAAATTATACCAAGCCTTTTAAAGCAAAAGCAAGCAGCCAC 415
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1166      ANAAAGAGATNANNNNNAANNTAAAAAAAANTAAANNAAAACANNAANAAAAAATTNNAN 1107
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY      416      AAATACAGTGAAGATCTGCAGTTAGATGCTGAACGACGCAATCATATGGACAA- 474
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1106      AANNAANAAAAAACAANNAANNAATATAAANNANNAANNAATNGAATATAAAAACAANNAAG 1047
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY      475      AGTCATGATGTACTTAACAAAGAGGTTAAACAGATGCTCTAAAGTTAAGACATCGACAC 534
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1046      ANTACAAANNAANAAAAAATAATANAANANANANNTATTAANANNAANNAACAGAAAA 987
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY      535      GACGCAGTGTTACTGGAAGCTACAGTGAATTAGATGATCCAAATATATCTACACGT 594
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      986      TNAANNAAGAAACAGACAGANNNANNAACAGANNACNNANNAANNAANNAANTTTAAAAA 927
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY      595      ATTCAGATGGAAAAAACCCATATAAGTATATGTCTAGTCTGGCAATATTCATTTT 654
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      926      AAGAANNANNAANAAAAAATAATAAATAANAGCANNAANANANANGANAGAAAAAAC 867
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY      655      AATCAGCAAAATTTATCAAGATGAATCAACACCACCAATTTGATATATCTGAAATCCAAAT 714
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      866      NAAAGAAAAAANNNNAANNAANNAANNAAAAAAATNNNAACATATGAANNAANNAANNAAC 807
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY      715      TTAACACACATCAACACATATTTGAATTTATTTACTTGAAGTCTTGATTTAACACAA 774
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      806      AANNAAAAAAATAAAGAGTAAATATNNNAGCAAAANNAAGATNNNAANNAATTTAA 747
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY      775      ATTCTCACATTTTATATAAAAAGAGACTTGGAA 808
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      746      AANAGANNAATNTTTNAAAGAAAAATTANNAA 713
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 14
CNS016EO
LOCUS      1201 bp      DNA      GSS      26-JUL-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC
            BACN1E09 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION  AL106652
VERSION     AL106652.1 GI:5623072
KEYWORDS
SOURCE
ORGANISM   fruit fly
            Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscophora; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE  1 (bases 1 to 1201)
            Genoscope.
            Direct Submission
            Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
            Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk/. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billard at CEPH (Centre
            d'etude du polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector

```



```

FEATURES                                pBelOBAC11.
Source                                  Location/Qualifiers
                                         1..1201
                                         /organism="Drosophila melanogaster"
                                         /plasmid="pBelOBAC11"
                                         /db_xref="taxon:7227"
                                         /clone_lib="DrosBAC"
                                         /clone="BACN15E09"
                                         /note="end : 17"

BASE COUNT      315 a       179 c       151 g       370 t       186 others
ORIGIN

Query Match          5.5%; Score 56.6; DB 121; Length 1201;
Best Local Similarity 40.8%; Pred.No. 0.0037;
Matches    95: Conservative   38: Mismatches 100: Indels    0: Gaps    0:

OY  563 TTGAATGATCATCCACAATATTACTACAGCTTTCAGAAGTGAAAAAACCTATAAAG 622
     |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db  943 TKTWTTKKGAMHTETKXMAAAAWMAARAARAAAATTTWWRASRDAAAAAAAAA 1002
     |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||

OY  623 TAATATGTCTAAGTCTGGCAATATTCATTTTAATCAGCAAATTTATCATCAAGATGAA 682
     |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db  1003 AAATWTTATWATWBRAAAAMTTTAAARTAAATATMATTAATWATTAATATATAT 1062
     |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||

OY  683 CACCAATTGGGATATATCTGCAAAATCCCAATTTTAACAAGCAATCAACACATATTGAAA 742
     |::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db  1063 WATWTTWTTTTAAAAAAMAAAAAAMTWTTAAAMTWMAAMWAATTTTWTWT 1122
     |::||::||::||::||::||::||::||::||::||::||::||::||::||::||

OY  743 TTATTTACTGAGCTCTGTGATTTACACACATCTTCACATTTATTAATAA 795
     |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db  1123 TTTTWTATTTWATWATWATWTTWTTAAAAATWATTTTATTTMTAAATATATATWTA 1175
     |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||

RESULT 15
LOCUS   CNS0039R/c
DEFINITION Drosophila melanogaster genome survey sequence TEI3 end of BAC #
        BACR08010 of RpCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL063932
VERSION   AL063932.1 GI:4941789
KEYWORDS  GSS.
SOURCE    fruit fly,
           Drosophila melanogaster
           Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
           Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
           Muscomorpha; Ephyroidae; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS   Genoscope.
TITLE      Direct Submission
JOURNAL    Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
COMMENT    BP 191 91106 EVRY cedex - FRANCE (E-mail : sequefegenoscope.cns.fr
           - Web : www.genoscope.cns.fr)
           Determination of this BAC-end sequence was carried out as part of a
           collaboration with the Berkeley Drosophila Genome Project (BDGP).
           The BDGP is constructing a physical map of the Drosophila
           melanogaster genome using these BACs. For further information
           please see http://www.fruitfly.org/The_BDGP_Drosophila_melanogaster_BAC_library.html
           The BDGP Drosophila melanogaster BAC library was prepared by Kazuo Osoegawa and
           Aaron Mammeter in Pieter de Jong's laboratory in the Department of
           Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
           NY. The library is named RpCI-98 and was constructed by partial
           EcoRI digestion of Drosophila DNA provided by the BDGP from the
           isogenic strain Y2: cn bw sp, the same strain used for the BDGP's
           pl and EST libraries. A more detailed description of the library
           and how to order individual BAC clones, the entire library, or
           filters for hybridization from the BACPAC Resource Center can be
           found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
           Location/Qualifiers
               1..1101
               /organism="Drosophila melanogaster"
               /db_xref="taxon:7227"

```

[illegible]

Search completed: November 19, 2000, 03:36:00
Job time: 12209 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 19, 2000, 04:23:24 ; Search time 2632.38 seconds
(without alignments)
2105.493 million cell updates/sec

Title: US-08-978-456-3
Perfect score: 1269
Sequence: 1 AACACCAATCCNATTGGGA.....GTTAACACAAGATCTATAG 1269

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1033670 seqs, 2183789903 residues

Total number of hits satisfying chosen parameters: 2067340

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pl1:*
10: gb_pr2:*
11: gb_pr3:*
12: em_fun:*
13: em_hum1:*
14: em_hum2:*
15: em_io:*
16: em_om:*
17: em_or:*
18: em_ov:*
19: em_pat:*
20: em_ph:*
21: em_pl:*
22: em_ro:*
23: em_sts:*
24: em_sy:*
25: em_un:*
26: em_v1:*
27: gb_htg1:*
28: gb_htg2:*
29: gb_in1:*
30: gb_in2:*
31: em_ba1:*
32: em_ba2:*
33: em_hum3:*
34: em_hum4:*
35: gb_pr4:*
36: gb_htg3:*
37: gb_htg4:*
38: gb_htg5:*
39: gb_htg6:*
40: gb_htg7:*
41: em_htg1:*
42: em_htg2:*
43: em_htg3:*

44: em_hum5:*
45: gb_pl3:*
46: gb_pr5:*
47: gb_htg8:*
48: gb_htg9:*
49: gb_htg10:*
50: gb_htg11:*
51: gb_htg12:*
52: gb_htg13:*
53: gb_htg14:*
54: gb_in3:*
55: gb_htg15:*
56: gb_htg16:*
57: gb_htg17:*
58: em_htg4:*
59: em_htg5:*
60: em_htg6:*
61: em_htg7:*
62: em_hum6:*
63: gb_htg18:*
64: gb_htg19:*
65: gb_da3:*
66: em_htg8:*
67: em_htg9:*
68: em_htg10:*
69: em_htg11:*
70: em_htg12:*
71: em_htg13:*
72: em_htg14:*
73: em_htg15:*
74: em_htg16:*
75: em_htg17:*
76: em_htg18:*
77: em_htg19:*
78: em_htg20:*
79: em_htg21:*
80: em_htg22:*
81: em_htg23:*
82: gb_pr6:*
83: gb_pr7:*
84: gb_htg20:*
85: gb_htg21:*
86: gb_htg22:*
87: gb_htg23:*
88: gb_ro:*
89: gb_sts1:*
90: gb_sts2:*
91: gb_sy:*
92: gb_un:*
93: gb_v11:*
94: gb_v12:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	352.2	27.8	3336	5	AR069579	AR069579 Sequence
2	289.2	22.8	3227	1	AF269646	AF269646 Staphyloc
3	289.2	22.8	3618	1	AF269345	AF269345 Staphyloc
4	288	22.7	3721	1	AF269926	AF269926 Staphyloc
5	176	13.9	2902	1	AF270130	AF270130 Staphyloc
6	103.6	8.2	6006	2	BSRIB	X51510 B. subtilis
7	103	8.1	28206	1	BACDIA	L09328 Bacillus su
8	103	8.1	218470	2	BSUB0013	Z99116 Bacillus su
9	100.8	7.9	4286	2	BARIBGENS	X95955 B.amyloliqu
10	98.8	7.8	18920	57	AC074319	AC074319 Staphyloc
11	84.8	6.7	14364	1	AE000675	AE000675 Aquifex a
12	69.4	5.5	18073	65	U32775	U32775 Haemophilus

C 13 68.4 5.4 11695 1 AE004298
14 67.6 5.3 2949 2 ECNUSB
15 67.6 5.3 11692 1 AE000148
16 67.6 5.3 139818 2 EC082664
17 64.4 5.1 2759 1 AF002857
18 64.4 5.1 4312 1 AF027202
19 63.8 5.0 14376 1 AE002531
20 62.6 4.9 11544 1 AE001343
21 60.6 4.8 349061 2 NMA222491
22 60.2 4.7 11404 1 AE003934
23 57 4.5 2307 2 BHE132928
24 56.4 4.4 105795 65 SYCCFNC
25 53.8 4.2 11456 1 AE002277
26 53.6 4.2 24887 2 BACSERA
27 52.6 4.1 3332 1 AF269888
28 52.6 4.1 170227 48 AF269712
29 52.4 4.1 170227 48 AC023138
30 52.2 4.1 110000 87 PFMA14P1_1
31 52 4.1 164353 37 AC011978
32 51.8 4.1 170413 64 AL159155
33 51.6 4.1 145670 10 AC008132
34 51.2 4.0 189841 64 AL139811
35 51 4.0 105567 35 AP001069
36 51 4.0 340000 35 AP001708
37 50.8 4.0 3542 45 YSCWTCG16
38 50.6 4.0 163226 27 AC004157
39 50.6 4.0 167712 84 AL356794
40 50.4 4.0 158482 50 AC025348
41 50.2 4.0 302250 83 U82671
42 50 3.9 1158 30 AF044857
43 50 3.9 184207 64 AL353748
44 49.8 3.9 1867 45 MTSCMJ23
45 49.8 3.9 12811 1 AE001819

ALIGNMENTS

RESULT 1
LOCUS AR069579 3336 bp DNA
DEFINITION Sequence 7 from patent US 5891672.
ACCESSION AR069579
VERSION AR069579.1 GI:7220467
KEYWORDS
SOURCE .
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 3336)
AUTHORS Wang,M., Ward,J.M., Warren,R.,Lloyd,Nicholas,R.,Oakley,
Palmer,L.,Marie, Pratt,J.M., Knowles,D.,JustinCharles,
Loretto,M.,Arthur, Mooney,J., Black,M.,Terence,
Burnham,M.,Karlsruhe, Debouch,C., Fedon,J.,Craig, Hodgson,J.,Edward,
Jaworski,D.,Dee, Reichard,R.,Winfield, Rosenberg,M., Trainl,C.,Michael
and Zhong,Y.Y.I.
TITLE Polynucleotides encoding GTP cyclohydrolase II (riba)
JOURNAL Patent: US 5891672-A 7 06-APR-1999;
FEATURES Location/Qualifiers
source 1..3336
BASE COUNT 1202 a 459 c 678 g 997 t
ORIGIN

Query Match 27.8%; Score 352.2; DB 5; Length 3336;
Best Local Similarity 88.4%; Pred. No. 8.6e-56;
Matches 473; Conservative 0; Mismatches 44; Indels 18; Gaps 8;

OY 734 ATGATATGCGATTCACATTCCTCAATATGTCACAGTCACAGTCCTTATACACCC 793
Db 1 ATGATATGCGATTCACATTCCTCAATATGTCACAGTCACAGTCCTTATACACCC 60
OY 794 GTGCGCGCTGTGTAGTAATGAAGTAGAGATTGTGTATGTGTGACACTTGAGAAA 853

Db 61 GTTGCGCGCTGTGTAGTAATGAAGTAGAGATTGTGTATGTGTGACACTTGAGAAA 120
OY 854 GGTGACAGCATCGCGAGCGTCAAGCACTTATGATGACACAAANATGCGAAGTGGC 913
Db 121 GGTGACAGCATCGCGAGCGTCAAGCACTTATGATGACACAAANATGCGAAGTGGC 180
OY 914 ACAGTTATATAGATTAGATAGAGCCATAGTATGTTGGTTACACCAACCCGTGTAAAC 973
Db 181 ACAGTTATATAGATTAGATAGAGCCATAGTATGTTGGTTACACCAACCCGTGTAAAC 240
OY 974 AAATATATGATTTGTAGATAGCANAAGTACTATTACCAACANAGCAATTCCTTAG 1033
Db 241 AAATATATGATTTGTAGATAGCANAAGTACTATTACCAACANAGCAATTCCTTAG 298
OY 1034 ACACACATTTGTTAGATAGACGTTACGGGCTCCACGCGTATTTGAGGGTGAATGCGTTG 1093
Db 299 ACACACATTTGTTAGATAGACGTTACGGGCTCCACGCGTATTTGAGGGTGAATGCGTTG 351
OY 1094 GATGATGACGCGCATCAATTTATACCAAGACTTTTAAAGCAAAAGCAAGCA 1153
Db 352 GATGATGACGCGCATCAATTTATACCAAGACTTTTAAAGCAAAAGCAAGCA 408
OY 1154 CTGCGCAAAATTTACAGTGAAGTCTTGAAGTTTACATGCGTAAACAAACGAATT 1213
Db 409 CT--GCCCAAAATTTACAGTGAAGTCTTGAAGTCTTGAAGTGAACAAACGCG--AAT 462
OY 1214 GATATGACAAAGTCATGATGATTTACTACAAAGGTTAAACAAAGATGTCTATA 1268
Db 463 GATATGACAAAGTCATGATGATTTACTACAAAGGTTAAACAAAGATGTCTATA 517

RESULT 2
LOCUS AF269646/c 3227 bp DNA BCF 01-AUG-2000
DEFINITION Staphylococcus epidermidis strain SRI clone step.1015d10 genomic
sequence.
ACCESSION AF269646
VERSION AF269646.1 GI:9623542
KEYWORDS
SOURCE .
ORGANISM Staphylococcus epidermidis.
REFERENCE Staphylococcus epidermidis.
Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Staphylococcus.
1 (bases 1 to 3227)
AUTHORS Kimmerly,W.J., Taylor,J.,David, Nelsen,A.J., Godlevski,M.M.,
Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
Listenbee,S., Ashanti,C., Altschuler,G., Mamo,L., Shepherd,N.S.,
Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
Furdon,P.J.
TITLE Transposon-mediated sequencing of the staphylococcus epidermidis
genome
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3227)
AUTHORS Taylor,J.,David, Kimmerly,W.J., Nelsen,A.J., Godlevski,M.M.,
Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I.,
Listenbee,S., Ashanti,C., Altschuler,G., Mamo,L., Shepherd,N.S.,
Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
Furdon,P.J.
TITLE Direct Submission
JOURNAL Submitted (22-MAY-2000) Departments of Genomic Sciences and
Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore
Drive, Research Triangle Park, North Carolina 27709-3398, USA
FEATURES Location/Qualifiers
source 1..3227
BASE COUNT 980 a 599 c 454 g 1194 t
ORIGIN

Query Match	Similarity	Score	289.2	DB 1	Length	3227
Best Local	Similarity 69.4%	Pred. No. 3.7e+44				
Matches	481	Conservative	0	Mismatches	177	Indels 41; Gaps
OY	346	AGATTATTAATTAATTAACCCGCTCATTTATTTTGTGAGAAACAATAGTATCATTT	405			
Db	2327	AAAATCTTTTAAAGTAGAGTATTAATTTGTATTTTAAATTTAAATAGTAGATTT	2268			
OY	406	TAAATGTAGTTGACACTACTAGTCAATAATCTATATACAAATTTTCAATATTAATCTT	465			
Db	2267	ACAATGTAGTAGAGAACTACTGAGCTCAATTAATC-ATACAAATTTCAATATAT--TTCTT	2211			
OY	466	TCGGGGCGAGGCGAATTCOCACCGGCGATTAATTAACCCGCGAGCTGCGAATTA--T	522			
Db	2210	TCGGGGCGAGGCGAATTCOCACCGGCGATTAATTAACCCGCGAGCTGCGAATTAATTTGT	2151			
OY	523	GTTTATATATAGTGGCTGATCTAGTAGAGATTTGTTAGAGCGGACAGTAAAGTGTGATGG	582			
Db	2150	TTTTATGTGTAGTGGCTGATCTAGTAGAGATTTGTTAGAGCGGACAGTAAAGTGTGATGG	2091			
OY	563	AGAAAGAAATGTTAAATTAATTCGACAAAGATTAATGTAGCGTATTTGTAAAA-----	630			
Db	2090	AGAAAGAAATGTTAAATTAATTCGACAAAGATTAATGTAGCGTATTTGTAAAA-----	2031			
OY	631	-----ATGCTGACAAATAGGCTTTATTAACATTAATTTTCTCCTTG	674			
Db	2030	ACTTATGTGCAATTTTAAAGTAGCTAATTAATGAGCTTATTAATTAATTAATAGATCTCTTAT	1971			
OY	675	CATTTAATTCATGATGTGAGGATTTTGTGTTATAGAGGTATCATTTAGTCAATTTA	734			
Db	1970	GCCTACATATATAGTATTAAGTTTTTTT-----GAGGTGATCATTTAGTATGATTTA	1918			
OY	735	TGGATTATGCGATTCGCACTTCCAAATATGTACAAAGTCAACAGGTGTTATCCACCG	794			
Db	1917	TGGATTATGCGATTCGCACTTCCAAATATGTACAAAGTCAACAGGTGTTATCCACCG	1858			
OY	795	TTGGCGGCTGTGTACTTATGAAGTAGAGATTTGTGTGTTGTGTGACACTTGAGAAAG	854			
Db	1857	TAGGATCCGTTGTGTGTAAAAACGGTAGGATTTGTAGGTGACACTTTAAAAAAG	1798			
OY	855	GTCGACACATGCGAGGTTCAAGCACTGATATGACACACANAAATGTTGAAGTGCA	914			
Db	1797	GAGATTAACATGCCAGACAGTACAAAGTATTTGAATGGCAGGTTTAAATGCCAAGTGCTA	1738			
OY	915	CGATTATATTAAGTTAGAGCCATGTAGTCAATTTGGTTCAACACACCCGTGTGTTACA	974			
Db	1737	CCATATAGCTTTCATTTAGAACCTTGACACACACCATGTTCAACACACCTGTGTGTCATA	1678			
OY	975	AAATTTATGATTTGTAAGATAGCANAAAGTAGTAT	1007			
Db	1677	AAATCATTTGAAAGGCGGACATATCTTAAGGTCATCT	1645			
RESULT 3	AF269345/c	3618 bp	DNA	BCT	03-AUG-2000	
LOCUS	AF269345					
DEFINITION	staphylococcus epidermidis strain sri clone step.1002d07 genomic					
ACCESSION	AF269345					
VERSION	AF269345.1	GI:9664666				
KEYWORDS						
SOURCE						
ORGANISM	staphylococcus epidermidis.					
REFERENCE	1 (bases 1 to 3618)					
AUTHORS	Kimmerly, W.J., Taylor, J. David, Nelsen, A.J., Godlewski, M.M., Rubino, M.A., Nelson, F.J., Rivers, P.R., Tortorella-Miller, L., Lisdenbee, S., Ashanti, C., Althuller, G., Mamou, L., Shepherd, N.S., Fuchs, R., Fleming, T., Guan, X., Du, L., Cain, D.H., Miller, G.S. and Fudon, P.J.					
TITLE	Transposon-mediated sequencing of the staphylococcus epidermidis genome					

JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 3618)
AUTHORS	Taylor, J. David, Kimmerly, W.J., Nelsen, A.J., Godlewski, M.M., Rubino, M.A., Nelson, F.J., Rivers, P.R., Torrella-Miller, I., Listenebee, S., Ashanti, C., Alshiller, G., Mamo, L., Shepherd, N.S., Fuchs, R., Fleming, T., Guan, X., Du, L., Cain, D.H., Miller, G.S. and Furdon, P.J.
TITLE	Direct Submission
JOURNAL	Submitted (19-MAY-2000) Departments of Genomic Sciences and Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore Drive, Research Triangle Park, North Carolina 27709-3398, USA
FEATURES	Location/Qualifiers
SOURCE	1. 3618 /organism="Staphylococcus epidermidis" /strain="SRI" /db_xref="taxon:1282" /clone="step.1002d07"
BASE COUNT	1099 a 709 c 518 g 1292 t
ORIGIN	
Query Match	22.8%; Score 289.2; DB 1; Length 3618;
Best Local Similarity	69.4%; Pred. No. 3.6e-44;
Matches 481; Conservative	0; Mismatches 171; Indels 41; Gaps 5;
OY 346	AGATTATTTTAAATTAACCGCTCATTTATTTGTTTTTACAAAACATATAGTATCATTT 405
Db 3110	AAAACTTTTATTAAGTAGTACCTATTTAAATTTGTTATTTTAAATTAATATGTAGATTT 3051
OY 406	TAAATGTAGTTGACATCTAGCTACTCAAAATACATAATCAATTTCAATATATTTCTT 465
Db 3050	ACAATGTGTGACAGACACTACGTCTCAATTAATC-ATAACAATTTCAATAT--TTCTT 2994
OY 466	TCGGGCGAGGGTGAATTTCCACCGCAGTAAATTAAGCTCGCGACCTGCTAATA--T 522
Db 2993	TCGGGCGAGGGTGAATTTCCACCGCAGTAAATTTGAGCTCGGACCCACTAATATGT 2934
OY 523	GTTTCATTTAGTGGCTATCTAGTGAATTCCTAGAGCCGACAGTTAAAGCTGTGATGG 582
Db 2933	TTTGTATTTAGTGGCTATCTAGTGAATTCCTAGAGCCGACAGTTAAAGCTGTGATGG 2874
OY 583	AGAAAGATGTTAATTTATCGACAAAGATTAATGTAGCGTATTTGTAAA----- 630
Db 2873	AGAAAGATGTTAATTTATCGACTTACACAATTTGCTGATGATGATAGCGTATTTGAT 2814
OY 631	-----ATGTGTACAAATAGGCTTATTTACGATAAATTTTCTCCTTG 674
Db 2813	ACTTATGTGCATTTTAAAGTATGCAATAGGCTTATTTAAAGATATAGATATCTCTTAT 2754
OY 675	CATCTTAATTCATGATGTGAAGATTTTTTTGTTATAGAGGTGATCATTTGATCAATTTA 734
Db 2753	GCCCTCAATTAAGTATAGGATTTT-----GAGGTGATCAATTTGATGATTTTA 2701
OY 735	TGGATTATGCGATTCACATCTCCAAATATGTCACAAGTCANACAGGTGTTATCCACCG 794
Db 2700	TGGATGATGCTATTCACATCTACCAAAAAGTAAATGAGCAACAGGTGTTATCCACCG 2641
OY 795	TTGGCGCTGTTGTATGATTAAGGTAGATTTGTTGATTTGTTGTCACACTTGAAGAAAG 854
Db 2640	TAGGATCCGTTGTTTAAAAACGGTAGATTTGATAGTTTATAGTTCACATTTAAAAAAG 2581
OY 855	GTGACACAGCATCGGAGGTTTAAGCACTTGATATGCAACAACANATGCTGAAGGTGCGA 914
Db 2580	GAGATTAACATGCGAAGTACAGCTATTTGAAGAGGCGGTTTAAATCCCAAGGTGCTA 2521
OY 915	CGATTATATTAACGTTNAGCGATGTAATCTTTTGGTTTCAACACCACCGTGTAAACA 974
Db 2520	CCATATAGCTTTCAATTAACCTTGACACACCATGTTTCAACACCACTTGTTGTCATA 2461
OY 975	AAATATGATTTGTAAGATGACANNAAGTAGAT 1007
Db 2460	AAATCATTTGAAGCGGCGCATATCTAAGGTCACT 2428

[illegible]

QY	735	TGCAATTAATGCATTAATCACTTCCAAATAATGTACAAGGTCAACAGAGTGTATATCCACCG	794
Db	258	TGATGATGATCAATTCACATACACAAANAATGTAATAATGACAAACAGGTGTATATCCACAG	199
QY	795	TTGGCGCTGTGTAGTAATGATGAGGTAGAGATTGTGGTATTGTGTGCACACTTTGAGAAAG	854
Db	198	TAGGATCCGTGTGTGTAAACCGTAGAGATTGTAGTTGAGTGTGCACATTTAAAAAAGG	139
QY	855	GTGCAACAGCATGCGGAGGTTCAGACACTGTATATGGCCACACANAATGCTGAGGTGGCA	914
Db	138	GAGATTAACATGCGGACACTACACACTTATGAAATGGCAGGTTTAATATCCAAAGTGTCTA	79
QY	915	CGAATTATATATACGTAGAGCCATGTAGTACTATTTGGTTGTTCAACACCACCTGTGTAAACA	974
Db	78	CCATATAGCTTTCAATTAGAACCTTGCACACACACATGGTTTCAACACCACCTGTGTGCATA	19
QY	975	AAATATTGA 984	
Db	18	AAATCATTTGA 9	
RESULT	5		
AF270130			
LOCUS	AF270130	2902 bp	DNA
DEFINITION	Staphylococcus epidermidis strain sri clone step.1051a10 genomic		
ACCESSION	AF270130		
VERSION	AF270130.1	GI:9624037	
KEYWORDS			
SOURCE			
ORGANISM	Staphylococcus epidermidis. Staphylococcus epidermidis Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group; Staphylococcus.		
REFERENCE	1 (bases 1 to 2902) Kimmerly,W.J., Taylor,J.David, Nelsen,A.J., Godlevski,M.M., Rudino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I., Listebeebe,S., Ashanti,C., Altschuller,G., Mamo,L., Shepherd,N.S., Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and Furdon,P.J.		
AUTHORS			
TITLE	Transposon-mediated sequencing of the Staphylococcus epidermidis		
genome			
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 2902) Taylor,J.David, Kimmerly,W.J., Nelsen,A.J., Godlevski,M.M., Rudino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,I., Listebeebe,S., Ashanti,C., Altschuller,G., Mamo,L., Shepherd,N.S., Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and Furdon,P.J.		
JOURNAL	Direct Submission		
FEATURES	Submitted (22-MAY-2000) Departments of Genomic Sciences and Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore Drive, Research Triangle Park, North Carolina 27709-3398, USA		
SOURCE	1. 2902 /organism="Staphylococcus epidermidis" /strain="SRI" /db_xref="taxon:1282" /clone="step.1051a10"		
BASE COUNT	1009 a 382 c 507 g 1004 t		
ORIGIN			
Query Match	13.9%: Score 176; DB 1; Length 2902;		
Best Local Similarity	68.2%: Pred. No. 2,9e-23;		
Matches 242; Conservative	0; Mismatches 113; Indels 0; Gaps 0;		
QY	653	AACGATAAATTTTCTCCTTGCATCTTAATTCATGATGTGAGGATTTTGTGTTAATAGA	712
Db	2139	AGCGATGACCTTTTACACGACATGATATGCGTCAACATATAGTATAGGTTTGTGGA	2198
QY	713	GCTGATCAATTTGATCAATTTATGATATGATCCGATTCATCTCCAAATATGTTACACAGT	772

[illegible]


```
/gene="lysa"
/citation=[1]
2496..2502
/gene="lysa"
/citation=[1]
2504..2510
/gene="lysa"
/citation=[1]
2666..2694
complement(2666..2694)
/note="ORFX19"
complement(2693..3565)
/note="ORFX19"
/codon_start=1
/transl_table=11
/protein_id="AA67474.1"
/db_xref="GI:410118"
/translation="MKIWIGMLAAVLLMPKYSLDAAVGVIVTLGADLSESDK
OKLDEMVPDNAVTVTNKEHEYLGYISNAGISRAISSITIAKSGSINVE
THNISGITDEMYLNLMTAGVADKAVYVAPEVSGTALTGLIKAYESSDEAISD
VKQVANQELVTTSELGDKIGNENAAALLAKIIEPRKNGVDPNKADIEQVDDASDL
NVTLDQKNDLVSLFNKMKKNADIDMGVSDLDKAKDKITIKIESDEKNTQKVID
FEVSIMNAIVSLFK"
/citation=[1]
3347..3353
/citation=[1]
3413..3419
/citation=[1]
3428..3434
/citation=[1]
3503..3509
/citation=[1]
3819..4290
/gene="ppib"
3819..4250
/gene="ppib"
/codon_start=1
/transl_table=11
/product="peptidyl-prolyl isomerase"
/protein_id="AA67475.1"
/db_xref="GI:410119"
/translation="MKTGYFLLEDGKNTIEFELYPEAAPGTANFEKLANEGFYDGLTF
HRVIGFVSQSGPHGTGTGPGYTICKTEGNPHHEGALSMHAGKDTGSGOFFI
VHEPQPLNGVITVFQKVTSGLEFAKKNMSGDMKVEVRGEG"
3982..3988
/gene="ppib"
/citation=[1]
4261..4290
/gene="ppib"
4287..4320
/note="inverted repeat"
5264..5367
/note="ORFX1"
/codon_start=1
/transl_table=11
/protein_id="AA67476.1"
/db_xref="GI:410120"
/translation="MOVLTLGLEOYGYAVLVLMELIALIGPEGFACVLYFOEKLN
AVLSMLAAGLGCSEHVVLDRV"
5445..5834
/note="ORFX2"
/codon_start=1
/transl_table=11
/protein_id="AA67477.1"
/db_xref="GI:410121"
/translation="MSYWIGYRGLGAPPAIOKAFASRGVLLAVIVLTFEFTVFGODRKI
ELSPLEFLVIGGELYEGIGIRLFHSLHPVHAPLESIVLVPPSEGLTAFVLFGCAY
LTVRSHKTRIQTVAFECILSYSSSG"
6022..6027
/gene="sips"
/note="putative"
6022..7323
/gene="sips"
6045..6051
```

```
/gene="sips"
/note="putative"
6683..6688
/gene="sips"
/note="putative"
6705..6710
/gene="sips"
/note="putative"
6740..7294
/gene="sips"
/codon_start=1
/transl_table=11
/product="signal peptidase"
/protein_id="AA67478.1"
/db_xref="GI:410122"
/translation="MKSBNVSKKSSILEMAKIYIAVLALLIRNFIFAPYVQDSM
IFILNKRERVNMTKITGEFDGSDIYLVGDDVHYKRIIGLPDIVEKKNQLYI
NGKRVDEPLYLANKRAKODGFHDLTDFGVKVPDNKTFVMDGRNRNSMDSRNGLGL
FTKQIDGTSFEFVPEFMKRTN"
7299..7323
/gene="sips"
7388..7393
/note="putative; ORFX4"
7410..7416
/note="putative; ORFX4"
7529..7873
/note="ORFX4"
/codon_start=1
/transl_table=11
/protein_id="AA67479.1"
/db_xref="GI:410123"
/translation="MGRIKRTITLLVLLLAGGYMYINDIELKDYPTAIGQTLSSSE
EEEYTIQEVKVTKIDGSEYHGAENGTIIIFNGKKLNQDLDIEKDGKIKAYFSKSR
IDGLIKVAKVND"
7872..7911
/note="ORFX4"
7937..7942
7960..7965
8126..8137
/note="putative"
```

```
Query Match      8.1%; Score 103; DB 1; Length 28206;
Best Local Similarity 60.6%; Pred. No. 6.3e-10;
Matches 166; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 729 AATTATGATTAATGCGATTCACCTTCCAAATATGTTACAAAGCTCANRACGTTGTAATC 788
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 8276 ATTATATGAAGCTGCGCTTAGACTTTCGCAAGCAGGCGCAAGACAGACGAATCCAAATC 8335

QY 789 CACCGCTGCGCGCTGTGTAGTTAATGAAGTAGAGATTGTGTATTGGTGCACTTGA 848
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 8336 CGCTCGTCGCGCTGTTCGTGTAAGAGCGACAAATTTGCGAATNGGCCCATTTA 8395

QY 849 GAAAGGTGACAAGCATGCGGAGGTTCAAGCACTTGATATGCGACAAANAATGCTGANG 908
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 8396 AATATGTGAAGCTCATGCGAAGTTCATGCCATCATATGCTGAGCAGCATGCAAGG 8455

QY 909 GTGCGACGATTTATATTAGCTTAGACCATGTAGTATTTGGTTCAACACCACTGTG 968
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 8456 GTGCGCAGATTATACCTTAGCACTGCAACCGTCGACGCAATTCAGCAAAAACCGCATGTG 8515

QY 969 TTAACAANAATTTATGATGTAGATGAGCANAAGT 1002
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 8516 CAGAAATTGATTATCAACTCTGATATCAAAAGAGT 8549

RESULT 8
BSUB0013/C BSUB0013 218470 bp DNA BCT 26-NOV-1997
LOCUS Bacillus subtilis complete genome (section 13 of 21): from 2395261
DEFINITION to 2613730.
ACCESSION 299116 AL009126
VERSION 299116.1 GI:2634723
```

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

Bacillus subtilis.
Bacillus subtilis.
Bacteria, Firmicutes; *Bacillus/Clostridium* group;
Bacillus/staphylococcus group; *Bacillus*.
1 (bases 1 to 218470)
Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G.,
Azevedo, V., Bertero, M.G., Bessières, P., Bolotin, A., Borchert, S.,
Boris, R., Boursier, L., Brans, A., Braun, M., Brinell, S.C.,
Bron, S., Brouillet, S., Brusch, C.V., Caldwell, B., Capuano, V.,
Carte, N.M., Choi, S.K., Codani, J.J., Connerton, I.F., Cummings, N.J.,
Daniel, R.A., Denizot, F., Devine, K.M., Dusterhoft, A., Ehrlich, S.D.,
Emmerson, P.T., Entian, K.D., Errington, J., Fabret, C., Ferrari, E.,
Foulger, D., Fritz, C., Fujita, M., Fujita, S., Fuma, S., Galizzi, A.,
Galleron, N., Gibi, S.Y., Glaser, P., Goffeau, A., Golightly, E.J.,
Grandi, G., Giuseppe, G., Guy, B.J., Haga, K., Halech, J., Harwood, C.R.,
Haupt, A., Hilbert, H., Holsappel, S., Hosono, S., Hullo, M.F.,
Ikaya, M., Jones, L., Joris, B., Karamata, D., Kasahara, Y.,
Klaer, Blanchard, M., Klein, C., Kobayashi, Y., Koetter, P.,
Koningstein, G., Krogh, S., Kumano, M., Kurita, K., Lapidus, A.,
Lardinois, S., Lauber, J., Lazarevic, V., Lee, S.M., Levine, A., Liu, H.,
Masuda, S., Mauel, C., Médigue, C., Medina, N., Melledo, R.P.,
Mizuno, M., Moestl, D., Nakai, S., Noback, M., Noone, D., O'Reilly, M.,
Ogawa, K., Ogihara, A., Oudega, B., Park, S.H., Paro, V., Pohl, T.M.,
Potterelle, D., Potvollik, S., Prescott, A.M., Prescan, E., Pujic, P.,
Punelle, B., Rapoport, G., Rey, M., Reynolds, S., Rieger, M.,
Rivolta, C., Rocha, E., Roche, B., Rose, M., Sadale, Y., Sato, T.,
Scanlan, E., Schleich, S., Schroeter, R., Scoffone, F., Sekiguchi, J.,
Sekowska, A., Seror, S.J., Serror, P., Shin, B.S., Soldo, B.,
Sotokin, A., Tacconi, E., Takagi, T., Takahashi, H., Takemaru, K.,
Takeuchi, M., Tamakoshi, A., Tanaka, T., Terasita, P., Tognoni, A.,
Tosato, V., Uchiyama, S., Vandenberg, M., Vannier, F., Vassarotti, A.,
Viari, A., Wambutt, R., Wedler, E., Wedler, H., Weitzenecker, T.,
Winters, P., Wiput, A., Yamamoto, H., Yamane, K., Yasumoto, K., Yata, K.,
Yoshida, K., Yoshikawa, H.F., Zumschein, E., Yoshikawa, H. and
Danchin, A.

The complete genome sequence of the gram-positive bacterium
Bacillus subtilis
Nature 390 (6657), 249-256 (1997)

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
JOURNAL

2 (bases 1 to 218470)
Kunst, F., Ogasawara, N., Yoshikawa, H. and Danchin, A.
Direct Submission
Submitted (18-NOV-1997) I. Moszer, A. Danchin, Institut Pasteur,
Regulation de l'Expression Genetique, 28 rue du Docteur Roux, 75724
Paris Cedex 15, FRANCE. E-mail: moszer@pasteur.fr,
adanchin@pasteur.fr Phone: +33 (0)1 45 68 84 41, Fax: +33 (0)1 45
68 89 48

FEATURES
source

Location/Qualifiers
1. .218470
/organism="Bacillus subtilis"
/strain="168"
/db_xref="taxon:1423"
complement(7. .18)
/gene="cmk"
complement(7. .696)
/gene="cmk"
complement(22. .696)
/gene="cmk"
/EC-number="2.7.4.14"
/note="alternate gene name: jofC, ypfC"
/codon_start=1
/transl_table=11
/product="cylidylate kinase"
/protein_id="CAB14221.1"
/db_xref="GI:2634724"
/db_xref="SWISS-PROT:P38493"
/translation="MEKKLSIAIDGPAAGSTVAKTIAEKKSYIYIDTGAMRYATY
AAIDENVDLDEKLAELIKRTDLEITTKGQGVVFNPGDVPFAIRPDESNVSA
AKHSVREEMKROQUGEGGVWMDGIDISTHLPVNAKIFLLASVEERAKRYEE
NVKKGFDVNYETLIEIARDKIDSERVSLRAADALEITDTLSIOEVADKILEA
VEOKSR"
complement(775. .951)

gene
terminator
CDS

CDS
/gene="ypfB"
complement(775. .951)
/gene="ypfB"
/function="unknown"
/note="alternate gene name: jofB"
/codon_start=1
/transl_table=11
/protein_id="CAB14222.1"
/db_xref="GI:2634725"
/db_xref="SWISS-PROT:P38492"
/translation="MKTFFRLIKLFLIOAILLGQFLFPHQIHIEPYSKYIQYIEGV
DKKEENRIETKH"
complement(96. .1649)
/gene="ypfA"
complement(96. .1649)
/gene="ypfA"
/function="unknown"
/note="alternate gene name: jofA"
/codon_start=1
/transl_table=11
/protein_id="CAB14223.1"
/db_xref="GI:2634726"
/db_xref="SWISS-PROT:P38491"
/translation="MIEGENVLEIIEENELKAKKAKVSTIENNELIAYVDVYTG
RTYILNDMEYVEFEGKDEVPYRISRIKGVKOKLQMCLEMPREKMKRIQROY
VRIDAVLDVQIOGNEEIRTLSTNSAGIIVLADGISTFQSGESTRLILPEBEH
TRQIEAVVRIRFNDPKSEKRMKLTLEYSEIAAGDQALLQYCIROLNKKRAME"
complement(1711. .3094)
/gene="ypdB"
complement(1711. .1733)
/gene="ypdB"
complement(1742. .3094)
/gene="ypdB"
/function="unknown"
/note="alternate gene name: jofB, yzua
similar to hypothetical proteins"
/codon_start=1
/transl_table=11
/protein_id="CAB14224.1"
/db_xref="GI:2634727"
/db_xref="SWISS-PROT:P38490"
/translation="MIRGILIAVLGAIYGTGYGCKEHOEKDAVLLHAENNYORAFH
ELTYQDQLDHDKGITLANNOSKSLSPALIDWRITSEAHNSVSQLPLMPNKTKE
LSKIDGFSYKTSRDLDOKPLDKNEVTSINKLYQOSEDIONELRHVQLVMSKNLKM
MDVEMLASDEKSDNTIINSEKTVKNYGASTGDLGSPSTKKEKGSILKKG
QISEQAKQIAERFAPDNDVSTIKVYSGKKNRDVYSIMKDPDKAAVYIMDTKKG
HPYVLQNEEVKQKISLNDGSRALAFIAKNGEFTDDLEIDESAGYDKIGVFSYVPV
ENKVRMYPEAIRKRVALDGEVVGESARDFLSHRKRTIPKPAITEAKSKLNKWO
VRETRLALITNELGQEVLCYEMLGITLNDTFRMYINAKGSEKVEKLNABPIYKDL
"
complement(3129. .4046)
/gene="sleB"
complement(3129. .4046)
/gene="sleB"
/note="alternate gene name: ypeA"
/codon_start=1
/transl_table=11
/product="spore cortex-lytic enzyme"
/protein_id="CAB14225.1"
/db_xref="GI:2634728"
/db_xref="SWISS-PROT:P50739"
/translation="MKSKGSIMACILIFFTITTFINFTETISAFSNQVYORGATGDV
VELQARLQNGVYNGKIDGVGMYGTWAVRNDODGLEGVGLVAGAKOTLLICKSK
YREYVMDOLNKNTFTTHYKQIDPKYQTPSKRAQOKRQOAEARQKPAEFTQKPK
ANANKQNNTPAKAKQDAVVAAMPQGFSSNDTRLIAQVYGEARGEPEIDGVAITAY
ILNRNLSPLPNSVAGVIEPLAFTVAGQITMOPNETAREAVLDAINGMPSEBAL
YVNPPTATSPWIMGRPOIRIKGILFCE"
complement(4162. .4841)
/gene="ypdC"
complement(4162. .4182)
/gene="ypdC"
complement(4185. .4841)
/gene="ypdC"

gene
terminator
CDS


```
|||||
Db 9385 ATCATATGTGGCTATGATGACATCTAGACGCCAGAC-TATAGTCTGATGGAG 9327
QY 585 AA 586
Db 9326 AA 9325

RESULT 11
AE000675/c 14364 bp DNA BCT 25-MAR-1998
LOCUS Aquifex aeolicus section 7 of 109 of the complete genome.
DEFINITION AE000675 AE000657
ACCESSION AE000675.1 GI:2982863
VERSION
KEYWORDS
SOURCE
ORGANISM Aquifex aeolicus.
Bacteria; Aquificales; Aquificaceae; Aquifex.
REFERENCE
AUTHORS Decker,G., Warren,P.V., Gaasterland,T., Young,W.G., Lenox,A.L.,
Graham,D.E., Overbeek,R., Snead,M.A., Keller,M., Anjay,M.,
Huber,R., Feldman,R.A., Short,J.M., Olson,G.J. and Swanson,R.V.
The complete genome of the hyperthermophilic bacterium Aquifex
aeolicus
JOURNAL Nature 392 (6674), 353-358 (1998)
MEDLINE 98196666
REFERENCE 2 (bases 1 to 14364)
AUTHORS Decker,G., Warren,P.V., Gaasterland,T., Young,W.G., Lenox,A.L.,
Graham,D.E., Overbeek,R., Snead,M.A., Keller,M., Anjay,M.,
Huber,R., Feldman,R.A., Short,J.M., Olson,G.J. and Swanson,R.V.
Direct Submission
JOURNAL Submitted (25-JUL-1997) Diversa Corporation, Genomics, San Diego,
CA 92131
COMMENT Putative indicates no similarity to known proteins
Hypothetical indicates similarity to a protein of unknown function.
FEATURES
source 1. 14364
/organism="Aquifex aeolicus"
/strain="vfs"
/db_xref="taxon:63363"
72. .410
/gene="rpsp"
/note="aq_123"
72. .410
/gene="rpsp"
/codon_start=1
/transl_table=11
/product="ribosomal protein S16"
/protein_id="AAC06492.1"
/db_xref="GI:2982872"
/translation="MAVRIRLAKFERKHHPIRIYVMDAKSPREKQYIDILCTYDPKR
KVLINVIPEKAEWLVKGVELSHRAKAILMNGILKEVVPGEYEMKRVGDYVFEKRE
SKRSKGEAA"
407. .646
/gene="aq_124a"
407. .646
/gene="aq_124a"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAC06503.1"
/db_xref="GI:2982883"
/translation="MSALKDIYELTAKELVDNKKVRYVTEIGEKTVVIELRVPDEL
KGVIGKGRIRIAALRTILTAIGRIKGRVLEILE"
646. .1044
/gene="aq_125"
646. .1044
/gene="aq_125"
/codon_start=1
/transl_table=11
/product="putative protein"
/protein_id="AAC06497.1"
/db_xref="GI:2982877"

gene
CDS
/translation="MKERLALALLSFFLGSGSFSEFYVAPTLFKLEBROQAGAVEKY
FPVYEGILIIIVGLISLFLGRSGKFLSLGILNLLLLLOEYIVIPKJLHGKATNYEL
FLKHGVSMGILNALILLTLGRKVLILIKR"
1049. .1492
/gene="aq_126"
1049. .1492
/gene="aq_126"
/codon_start=1
/transl_table=11
/product="putative protein"
/protein_id="AAC06498.1"
/db_xref="GI:2982878"
/translation="MNKEGVKELISYGNMIALEPIINTKLEIKLKAQIITLS
LDVCKTEKEMKMDLEGVNEVAEIAKEMTONRMERMMLFEPSSEVLKERMEG
DIYEDKHELSTLCKISKILKRLDLKRNKGILKIRRLIFR"
1492. .2310
/gene="aq_127"
1492. .2310
/gene="aq_127"
/codon_start=1
/transl_table=11
/product="putative protein"
/protein_id="AAC06499.1"
/db_xref="GI:2982879"
/translation="MLSELYVYLAMAFQAFTEGAGIFVIGILSLEPEKTVIYSS
VYNLIGIISMLFIKIRPNPKILFPIAGSVAGIGYSAKLMEIDREVLKALIGF
VFLGITYDFLYQTSNLSKRLKASPPVCGFTGFLSGIFGLIGMGPPPVYLVNOYCRD
LNTKTLISFFPATNVFRILFYLDGTEWSELIPALGLGAVGAFAGIYSRYF
TPYFKRPLISLVLGLFLEIGKEPSHDFHSHVIEESPYGGLSKSHKPGN"
complement(2134. .3552)
/gene="aq_128"
complement(2134. .3552)
/gene="aq_128"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAC06502.1"
/db_xref="GI:2982882"
/translation="MORITVNPNEPYLSVIRKVKLSPIIVVNLITVENMISILY
SSISSAVAATGFSLSLMTFYSLSLAUSYSGTINLIAQVGAKKDPSTILNGJLSF
LISLPLEFYGRDFVLEFMKVLGASETVSLAKEVLTPLFMPPIGIFLNTFYGYANGA
GDTKTPMKVAILIMNLTHIGATYTLINGFGFPKKGVEBAGWGAISSELAFFIYFLL
IEFKKPEPLHLREPKILFKMVRGLGTALRRATITTSFNVEFGVLAQKVLAAHQ
IGLRESISPMWIGFVMASTTLAGONYGARNYGMVAVTSAHFTLVNLSGLIL
ILPHYLVYPPSRDPEVIEMASYYLQIYGISQPMANASISGALKMGKTHIPLFVN
ISSFWLFRILIPSYFLKLVHSPLVPWGMFETAVRALFYTFKRVGKLLQTFQKE
EEPEGDEAGDCKPLDEVRSKYT"
3567. .4031
/gene="r1bh"
3567. .4031
/gene="r1bh"
/codon_start=1
/transl_table=11
/product="riboflavin synthase beta subunit"
/protein_id="AAC06489.1"
/db_xref="GI:2982869"
/translation="MQIYEGKLTAEGLRFGIVASRFNHALVDRLVEGAIDCIVBGR
BEDITLVAVPGSWEIPVAVAGELARKEDIDAVIAIGVLRGATPHFDYIASVSGLAN
LSLELRKPIITPGVITADTLFOAIRAGTKHNGKGMALSAIEMANLFKSLR"
4028. .4474
/gene="nusb"
4028. .4474
/note="aq_133"
/gene="nusb"
/codon_start=1
/transl_table=11
/product="transcription termination Nusb"
/protein_id="AAC06491.1"
/db_xref="GI:2982871"
/translation="MRYRKGAARDPAFLVLYRMDLRGENPGELFKVEVEEKNIKNKDAY
EYAKIVDAVRIIEIDISITEKLGKSIDRLGTVENRNALRGVAELIFLKSPEGR
VFIDIVDKKVADEKAGKFPVGLSLAIYKAYITSSNEKSLKSE"
```

[illegible]

QY	789	CACCCGTTGGCCCTTGTAGTATGAAAGGAGGATTTGGTATGTGTGACACTTGA	848
QY	8150	CTACGCTGGGACCGCTTGTCTGTAAGGAAGGAAAAATCTGGGATTAGTATTCACGAA	8091
QY	849	GAAGAAGTGACAGCATCGGAGAGTTCAAGCACTTGATATGACACAGAMATGCTGAAG	908
Db	8090	AAGCAGGAAGAGCTCAGCCTGAAGTATGCGTCTCGTCAGGCGGGGAAAAAGCAAAAG	8031
QY	909	GTGCGACGATTTATATATACGTTTAGAGCCATGTAGTCATTTTGGTTCAACACCCTGTG	968
Db	8030	GGGCAACCTTTACCTTACCTTACCCTTTAGGCCGCTGACTACCTTCGGGAACCTCCTCGA	7971
QY	969	TTAACAAATTTTGTATTTAGTATGTAAGAGCAGMAGTAGTATTTACNCAACANAGACATTC	1028
Db	7970	CGGAGCGGATTAATAGAGCGGGGATTTAAAGGGTTGTAGTGCACCGCTTGACCCAAATC	7911
QY	1029	GTTAGACACACATGGGTGATGAGAGCTTACGGGGCTCCACGGTATTTAGSGTTGAATG	1088
Db	7910	CTTTGATGCTCTGGAAAGGTGTTGAAAACTGAGAAATGGCGGATTAGAAGTGATGTCG	7851
QY	1089	CGTTGATGATGACAGCGGCATCACTAATATACCAAGAAGCTTTTAAAGCAAAAGCAA	1148
Db	7850	GTGTATGCGAAGAGGAGCGGCACAGAACCAAGAGATTTTCTTACTTACATCACTCAGG	7791
QY	1149	AGCACTTGCACCAAAATTACAGT	1173
Db	7790	AAAGACTTACATACCTTAATG	7766
RESULT	12		
LOCUS	U32775	18073 bp	DNA
DEFINITION	Haemophilus influenzae Rd section 90 of 163 of the complete genome.		
ACCESSION	U32775	L4203	
VERSION	U32775.1	GI:1573951	
KEYWORDS			
SOURCE			
ORGANISM	Haemophilus influenzae Rd. Haemophilus influenzae Rd. Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae; Haemophilus.		
REFERENCE	1 (bases 1 to 18073)		
AUTHORS	Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerevage, A.R., Bolt, C.J., Tomb, J., Dougherty, B.A., Merrick, J.M., McKenney, K., Sutton, G.G., FitzHugh, W., Fields, C.A., Gocayne, J.D., Scott, J.D., Shirley, R., Liu, L.I., Glodek, A., Kelley, J.M., Weidman, J.F., Phillips, C.A., Spriggs, T., Hedblom, E., Cotton, R.C., Utterback, T., Hanna, M.C., Nguyen, D.T., Saudek, D.M., Brandon, M.C., Fine, L.D., Fitcham, J.L., Fuhman, J.L., Geoghegan, N.S., Gnehm, C.L., McDonald, L.A., Small, K.V., Fraser, C.M., Smith, H.O. and Venter, J.C.		
TITLE	Whole genome random sequencing and assembly of Haemophilus influenzae Rd		
JOURNAL	Science 269 (5223), 496-512 (1995)		
MEDLINE	95350630		
REFERENCE	2 (bases 1 to 18073)		
AUTHORS	Tatusov, R.L., Mushegian, A.R., Bork, P., Brown, N.P., Hayes, W.S., Borodovsky, M., Rudd, K.E. and Koonin, E.V.		
TITLE	Metabolism and evolution of Haemophilus influenzae deduced from a whole-genome comparison with Escherichia coli		
JOURNAL	Curr. Biol. 6 (3), 279-291 (1996)		
MEDLINE	96389784		
REFERENCE	3 (bases 1 to 18073)		
AUTHORS	White, O., Clayton, R.A., Kerevage, A.R. and Fleischmann, R.D.		
TITLE	Direct Submission		
JOURNAL	Submitted (25-JUN-1995) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA		
REFERENCE	4 (bases 1 to 18073)		
AUTHORS	White, O., Clayton, R.A., Kerevage, A.R. and Fleischmann, R.D.		
TITLE	Direct Submission		
JOURNAL	Submitted (27-SEP-1997) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA		
REMARK	The H. influenzae sequence has been updated by R. Fleischmann. New database matches have been assigned, product names have been		

Improved, and a number of frame shifts have been corrected. We gratefully acknowledge the work of Tatusov et. al. We have incorporated their annotation into the /notes fields of the corresponding H. influenzae genes

5 (bases 1 to 18073)

White, O., Clayton, R.A., Kerlavage, A.R., Fleischmann, R.D., Peterson, J., Hickey, E., Dodson, R. and Gwinn, M.

AUTHORS

TITLE

Submitted (28-MAY-1998) The Institute for Genomic Research, 9712

JOURNAL

Medical Center Dr, Rockville, MD 20850, USA

REMARK

The whole genome was shifted by 588 nucleotides for a new start

COMMENT

On Sep 30, 1996 this sequence version replaced gi:1221677.

FEATURES

source

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

/codon_start=1
/transl_table=11
/product="enolase (eno)"
/protein_id="AAC2590.1"
/db_xref="GI:1573953"

/translation="MAKIVKIGREIISDRGNPTVEAEVHLXGCEVGLAABSGASTG

SRFALEIRGDDSRFLGKGVLAANAANNIEAOATVGDATNOAEDIDIMIDGTEH

KNFGANILAVSIANAKAAASGLPLVAYIAELNGTAGVSMPLPMNIIINGEHA

DNNVDIOEFMIOPVGAKEITREBALRIGAEVFNHIAKAKKAGSTAVGDEGAPNLAS

NADALCIKENAEKAGYVUGKDYTLAMDACASSEFYKKEKNGEMKEKGSFTSOETH

YLEELCKEPIVSIEDGODESWGEFAVOTKVGVGLVGLDLEFVNTKILKEGLEK

GIANSILIFENOGSLTERLAIKAKMDAGYAVISHRSGEEDAIADLADVATAGQ

IKTGSMSRSDRIAKYNQLIRIEALERACCTPAFPGLKAVKQGA"

/complement(3855..5060)

/gene="H10933"

/complement(3855..5060)

/note="similar to SP:P37631 PID:912464 GB:U00096

PID:2367234 percent identity: 58.35; identified by

sequence similarity; putative"

/codon_start=1

/transl_table=11

/product="conserved hypothetical transmembrane protein"

/protein_id="AAC2591.1"

/db_xref="GI:1573954"

/translation="MSOYSENIHIIIGAGAGLFCQAOLAKIGKSVTFEDNGKIKRIL

MSGGRCNFTNLEVPRAHYLSOHPRVKSAIARVTMDPISLVAEGITVYHREKGL

FCDBGEQIVEMKSECDKYGAKILRESEVERIONDEKRAFLVQVASTONOCKNL

IVATGLSMPLGATAPFGVQIAEFGFIPVTPASVPTVETEDLFLAALSGISLPV

ITVALGSKFENQDLTFHRGISGPAVALQSNWQPSVEIDLLPHNVEEIDNOAKQ

SSPKOKLITLVLPLKLVLELIEQIVODEVIANISVRKYNLVDFIHHMEFTPG

TEGYRAEYTLGAGVDVTKVISTMESNQVSGYFIEGIVLDVYGMILGYNFQWAMSSAY

ACALSSRSQ"

/complement(5057..6211)

/gene="H10934"

/complement(5057..6211)

/note="similar to SP:P32711 GB:X72298 PID:396410

PID:404307 GB:U00096 percent identity: 50.41; identified

by sequence similarity; putative"

/codon_start=1

/transl_table=11

/product="cytochrome C-type biogenesis protein (nrfp)"

/protein_id="AAC2592.1"

/db_xref="GI:1573955"

/translation="MKSIDITFRGFIQVQKVFYFATLLPAFSLFTQAEWDTYFQ

NDDDRRAVELAKSLRCPQONONLVESPLAVDYRLRVKMDGSKNOOITDKMT

AREGDTVNRKPPFKMTALMLPVALLLIAVLLFYSNRKQFSKRVAAQLENDIEI

ISLPSTFGSGSPKQSGSKGKVSRTYFVLLIIPATYIFSLDRFSVOGE

OSMTIEGHONVENDEKHNENYERKONKLRTPDPAEETMLGAEKAYONNEFDSALV

AYSNAEKLSGSKRNILGLAATLALYQAQOMQKVLOLNEALAKKNEYSLSLAT

IALENRQYQAGNYLQOQLDLSGAAVDRSRIORMKMLDFLRGEGQNP"

/complement(6208..6738)

/gene="H10935"

/complement(6208..6738)

/note="similar to SP:P33926 PID:405920 GB:U00096

PID:178523 percent identity: 42.60; identified by

sequence similarity; putative"

/codon_start=1

/transl_table=11

/product="thiol:disulfide interchange protein (dsbe)"

/protein_id="AAC2593.1"

/db_xref="GI:1573956"

/translation="MKKKLIFFTPLEFVLLGVCILLIAGINQDPKRIASALLDKPVPEF

YQANLHPSQIVSPKEPKRQPLFLNVMGSCGCKEHLHLIEIATLPIVAVDRDN

RONGIKMLKRMGNPFILITIDSHGELGKLGVGABETVLDENGVIARRHSGILDK

TYQTVFLPKIEALXNK"

/complement(6738..8645)

/gene="H10936"

/complement(6738..8645)

/note="similar to SP:P32710 GB:X72298 PID:581146 GB:U00096

PID:1790511 percent identity: 42.42; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="cytochrome C-type biogenesis"
/protein_id="AAC23594.1"
/db_xref="GI:1573957"
/translation="MLPELGFLLLLATASAPFLAVPOFGLFKKPTLINAAMPLSY
LFTLATSTLSILASFAVDPELLEVAAHSNOLPFEKVAATGHEGSMLEPLFS
LSMLAFAFENRKRDRFSSNOSLSLLICGFAVFLFISNPGRIFFPAPABEGRDL
NPMLODVLFIHFPLLYGVYGFVAFNEMASLSALTYNSARQIARSMGWLVSWFL
TIGVIGAMMAVYELGNGMWFMDVENASLMPWLGLALHSLMATEKQGVFSYWT

Query Match 5.5% Score 69.4; EB 65; Length 18073;
Best Local Similarity 49.7% Pred No. 0.0011;
Matches 172; Conservative 0; Mismatches 174; Indels 0; Gaps 0;

QY 657 ATAAATTTTCTCCTTGCATCTTAATCATGATGAGGATTTTGTATAGAGTG 716
DB 15823 ATGATCAATTTACTTCGAAATGAAAAATTAAAGATTATATGCTGATTTTCT 15882
QY 717 ATCATTTGATCAATTATGATTTATGCAATTCCTCCAATATGACAAAGTCANA 776
DB 15883 CACAAGATTGCGTATTATGCAACGTCCTTGAATTTGGCAGCCAAAGGTCAATATACCA 15942
QY 777 CAGGTGTTATCCACCCGTTGGCGTGTGTAGTTATGAGGATGATTTGATTTG 836
DB 15943 CTACACCCGAACCATGCGTGGGATGCGTATGTTAAATGTTGAATTTGGGGAGG 16002
QY 837 GTGCACACTTGAAGAAAGTGACMACATGCGAGGTTCAAGCACTGATATGACACAAC 896
DB 16003 GCTTTCATTTTAAAGCAGACACCTCATGCTGAACGCGTGGCTAGCTCAAGCTGTG 16062
QY 897 AANAATCTAAGTGGCAGATTTATATACGTTAAGCCATCTACTATTTGGTTCA 956
DB 16063 AAAATGCTAAGAGACACCGCTTACGTCACCTTAGAGCTTGCGCTCATTTAGTGCGA 16122
QY 957 CACCAACCCGTTGTTACAAATATATTGTTAGATACANAGT 1002
DB 16123 CACCACTTGTGCTATGAGATTATGAGGCGGCGCTAGTGAAGT 16168

RESULT 13
AE004298/c 11695 bp DNA BCT 31-JUL-2000
LOCUS Vibrio cholerae chromosome I, section 206 of 251 of the complete
DEFINITION Chromosome.
ACCESSION AE004298 AE003852
VERSION AE004298.1 GI:9656821
KEYWORDS
SOURCE
ORGANISM
Vibrio cholerae.
Vibrio cholerae.
Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.

REFERENCE 1 (bases 1 to 11695)
Heidelberg, J.F., Eisen, J.A., Nelson, W.C., Clayton, R.A., Gwin, M.L.,
Dodson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Umayam, L.A.,
Gill, S.R., Nelson, K.E., Read, T.D., Tettelin, H., Richardson, D.,
Ernstlaeva, M.D., Vamathevan, J., Bass, S., Qin, H., Dragoi, I.,
Sellers, P., McDonald, L., Uitterback, T., Fleischmann, R.D.,
Nierman, W.C., White, O., Salzberg, S.L., Smith, H.O., Colwell, R.R.,
Mekalanos, J.J., Venter, J.C. and Fraser, C.M.
DNA Sequence of both chromosomes of the cholera pathogen Vibrio
cholerae

TITLE Nature 406, 477-483 (2000)
2 (bases 1 to 11695)

JOURNAL Heidelberg, J.F., Eisen, J.A., Nelson, W.C., Clayton, R.A., Gwin, M.L.,
AUTHORS Dodson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Umayam, L.A.,
Gill, S.R., Nelson, K.E., Read, T.D., Tettelin, H., Richardson, D.,
Ernstlaeva, M.D., Vamathevan, J., Bass, S., Qin, H., Dragoi, I.,
Sellers, P., McDonald, L., Uitterback, T., Fleischmann, R.D.,
Nierman, W.C., White, O., Salzberg, S.L., Smith, H.O., Colwell, R.R.,
Mekalanos, J.J., Venter, J.C. and Fraser, C.M.
Direct Submission

TITLE

JOURNAL
FEATURES
SOURCE

Submitted (14-JUN-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
Location/Qualifiers
1. 11695
/organism="Vibrio cholerae"
/strain="N16961"
/serotype="O1"
/db_xref="taxon:666"
/note="biotype: E1 Tor"
complement(97. 825)
/gene="VC2260"
complement(97. 825)
/note="similar to GB:D26562 SP:P07906 GB:M5106 PID:146727
PID:473823; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="ribosomal protein s2"
/protein_id="AAP95404.1"
/db_xref="GI:9656822"
/translation="MASVSMRDMLTAGVHFHGHOTRYNPKMKOIFGARBRVHINLE
KTYPMENBALAEALAKVEKKGYLFVSTKRAASGVKALASNOYVYNNRWIGMT
NMKTVRSIRKLKELEVOSTDGTFDKLTRKEALRKREMERLESIGEDIKOMGLPDA
LEVIDADHEHIAIKENANNIGIPFAVVDITNSPDGVYIIPGDDAIRAOLYINAA
QAIKEGKRNKDVAAVAEKDCEVEAE"
1166. 2008
/gene="VC2261"
1166. 2008
/gene="VC2261"
/note="similar to GB:D26562 SP:P07906 GB:M5106 PID:146727
PID:473823; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="methionine aminopeptidase"
/protein_id="AAP95405.1"
/db_xref="GI:9656823"
/translation="MSIKIRNAVEIEKRYAAGRLAELEMEIEPHVAKAYTTELDQI
CHKYITEVOGAIAPALNYHGFPKSICTSIINHICVCPASEDETFQIOPPAVLRDD
LTINDIVIKDYGDTSKMELIGDVAIEDIRKCHVROECTYALAKOVKGVOLGIG
TTTEKHITKNNKNNPRKFSIVRDYCHGGAEGAEHREPOVYHKNRSRYLRBSMIFT
IEPWINAGKFCGRIDDESMVTYATDCKSAQWETHLLVATGCEIITLSEESLPRI
LNAA"
2098. 4743
/gene="VC2262"
2098. 4743
/gene="VC2262"
/note="similar to GB:M96431 SP:P27249 PID:146155 PID:49393
PID:473822; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="protein-P-II uridylyltransferase"
/protein_id="AAP95406.1"
/db_xref="GI:9656824"
/translation="IMGRAMPYOSPIFPOEPLTVESLKQOLESFTEYOKOEFDHHP
VTVDLGRSEMDLLRLMWFEGFDELVEISLVAVGSGEGELHPSIDIDLVSQ
PLSFQVANKISQFLTLMDLKLEIGHAVRYEQAELGKADLVATLMQARILGCE
ETPRRLKVVHISESPSEIFQYOKAVREKORAKYRDYTNNEPDIKSPGIGRDR
TLSVARRHGATSLYEMSRGFTLDAEYRELVESCOPILRVRRALITELKRYNRT
FAHVOYVARHIGYGBGNRGLEEMMKKEFTLRVALELKNMLITFKALINNGEAE
VAIIDDFORGNMIEARKPALFQARPEITLDFLHASDSTIESVAPATRNQIAR
RLNKFPLHTLPAREKFIELVRNHNLAKAFSQMHKGLVLAAYLPQNOIVGOMQPL
FHVYTVDSHSLRLKHTLPSDANNRHRPCCREIYPRICKELLILAAIFHDGGR
GGDSSEIGADAPDCEIEHGISKPEAKLVAMVYNNHLSMVTQGRDIDPDVIEFA
KRIVDERLETIVLVDICATPELINSKRRLLAELEYSTORALRGLENVDRA
KRIHNOOMSALLRKGEFSRELEVMORKAYFLRHHTKQIAMWCTHLRREDS
KPLVLSEKATRGEEVFIYTKDQAAFAFVAALVADLNNDNHOAQIMASDGVYLT
EMVLDNGOALAEEDRHQALIRHLYHVEDSDPTQKARRIPRNQSHKVTQVDFLT
KSKRTIMEFVALDPLGLATVATGAFVLELNDHAAKITTTGERAEDELITLNAQGR
LNEEDQHLRKLEHVAELAPSQ"
complement(4818. 4916)
/gene="VC2263"
complement(4818. 4916)

gene
CDS

gene
CDS
/gene="VC2263"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF95407.1"
/db_xref="GI:9656825"
/translation="MSMLGTHSLAAYLQLOLWVYFTTKALNL"
4989..5402
/gene="VC2264"
4989..5402
/gene="VC2264"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF95408.1"
/db_xref="GI:9656826"
/translation="MRVILIKRLPMPFPLVGLGNEPQIRISLROBAHKDYKITF
KKQKELSLASKVEKYPRQIKNVLDSSGSHQYKEVTEINRLTLVIDELNKIKPEWM
GEVDVKORILNDLRLEKVVASKIAIEDADLOKLN"
complement(5443..5952)
/gene="VC2265"
complement(5443..5952)
/gene="VC2265"
/note="similar to GB:U00096 PID:1773102 PID:1786620
PID:2329841; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="phosphatidylylglycerophosphatase A"
/protein_id="AAF95409.1"
/db_xref="GI:9656827"
/translation="MSDPRSNLSLSNPWHLATGFGSGSPIDVGTMTLASVFFYLT
LAQLEWIIAYVAVVAFSLIGIKCOVTSDDQVNDHGSIVDEEAGFWITWLYPLQX
LPVPMKMLVGFVLEFRFMDLKPWIPMLDKRVHGGGIMIDDLVAGVMAISLAI
GYWLEMLV"
complement(5949..6953)
/gene="VC2266"
complement(5949..6953)
/gene="VC2266"
/note="similar to GB:U00096 SP:P77785 PID:1773101
PID:1786619 PID:2329840; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="thiamin-monophosphate kinase"
/protein_id="AAF95410.1"
/db_xref="GI:9656828"
/translation="MQRNQRVMMWFGEFNILIDKYSNRQAQRKDVHLAAGDCAIVK
VPENSRVASIDTVAAGTHFLAQNPAWVAHKALASNTSLAAGATPAWVSLTLTP
EIDDAWLPFGCAFEFLAKYVNVOLIGDTRKGPSITLVYGFLEPQEARLRSAGVLA
GDMVITGDGDSQGLDVILIDPEKRHLIPFADIDLEORHYSITPVIQALVHLASSA
IDISGLAADIOLHILRRSNVGCASIDVSLPLSKLELOVDSVTSAQOYLITGSEYEL
CFTIPEBRKGSERNLAHCGTRKVTICGQIRPAGITFELNQGKLLDMQDLAIGTDHKVKA
"
complement(7014..7484)
/gene="VC2267"
complement(7014..7484)
/gene="VC2267"
/note="similar to GB:M26839 SP:P04381 GB:X00681 GB:X64395
PID:146976; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="N utilization substance protein B"
/protein_id="AAF95411.1"
/db_xref="GI:9656829"
/translation="WGASVKKPAARRNAFOALQAIYSQITKENVATIEOFLTSGKY
DEEERRAEPALAAIEDVSYFRDLACGVVNHHELSKLPFVSRMODIMNELAL
LRLAYEMTRREDVYKVINAEATELAKVFAEDSKFVNVGLDKAAPVHKKA"
complement(7484..8005)
/gene="VC2268"
complement(7484..8005)
/gene="VC2268"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="phosphatidylylglycerophosphatase A"
/protein_id="AAF95412.1"
/db_xref="GI:9656830"
/translation="MSDPRSNLSLSNPWHLATGFGSGSPIDVGTMTLASVFFYLT
LAQLEWIIAYVAVVAFSLIGIKCOVTSDDQVNDHGSIVDEEAGFWITWLYPLQX
LPVPMKMLVGFVLEFRFMDLKPWIPMLDKRVHGGGIMIDDLVAGVMAISLAI
GYWLEMLV"
complement(5949..6953)
/gene="VC2269"
complement(5949..6953)
/gene="VC2269"
/note="similar to GB:U00096 SP:P77785 PID:1773101
PID:1786619 PID:2329840; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="thiamin-monophosphate kinase"
/protein_id="AAF95413.1"
/db_xref="GI:9656831"
/translation="MQRNQRVMMWFGEFNILIDKYSNRQAQRKDVHLAAGDCAIVK
VPENSRVASIDTVAAGTHFLAQNPAWVAHKALASNTSLAAGATPAWVSLTLTP
EIDDAWLPFGCAFEFLAKYVNVOLIGDTRKGPSITLVYGFLEPQEARLRSAGVLA
GDMVITGDGDSQGLDVILIDPEKRHLIPFADIDLEORHYSITPVIQALVHLASSA
IDISGLAADIOLHILRRSNVGCASIDVSLPLSKLELOVDSVTSAQOYLITGSEYEL
CFTIPEBRKGSERNLAHCGTRKVTICGQIRPAGITFELNQGKLLDMQDLAIGTDHKVKA
"
complement(7014..7484)
/gene="VC2270"
complement(7014..7484)
/gene="VC2270"
/note="similar to GB:M26839 SP:P04381 GB:X00681 GB:X64395
PID:146976; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="N utilization substance protein B"
/protein_id="AAF95414.1"
/db_xref="GI:9656832"
/translation="WGASVKKPAARRNAFOALQAIYSQITKENVATIEOFLTSGKY
DEEERRAEPALAAIEDVSYFRDLACGVVNHHELSKLPFVSRMODIMNELAL
LRLAYEMTRREDVYKVINAEATELAKVFAEDSKFVNVGLDKAAPVHKKA"
complement(7484..8005)
/gene="VC2271"
complement(7484..8005)
/gene="VC2271"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="phosphatidylylglycerophosphatase A"
/protein_id="AAF95415.1"
/db_xref="GI:9656833"
/translation="MSDPRSNLSLSNPWHLATGFGSGSPIDVGTMTLASVFFYLT
LAQLEWIIAYVAVVAFSLIGIKCOVTSDDQVNDHGSIVDEEAGFWITWLYPLQX
LPVPMKMLVGFVLEFRFMDLKPWIPMLDKRVHGGGIMIDDLVAGVMAISLAI
GYWLEMLV"
complement(5949..6953)
/gene="VC2272"
complement(5949..6953)
/gene="VC2272"
/note="similar to GB:U00096 SP:P77785 PID:1773101
PID:1786619 PID:2329840; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="thiamin-monophosphate kinase"
/protein_id="AAF95416.1"
/db_xref="GI:9656834"
/translation="MQRNQRVMMWFGEFNILIDKYSNRQAQRKDVHLAAGDCAIVK
VPENSRVASIDTVAAGTHFLAQNPAWVAHKALASNTSLAAGATPAWVSLTLTP
EIDDAWLPFGCAFEFLAKYVNVOLIGDTRKGPSITLVYGFLEPQEARLRSAGVLA
GDMVITGDGDSQGLDVILIDPEKRHLIPFADIDLEORHYSITPVIQALVHLASSA
IDISGLAADIOLHILRRSNVGCASIDVSLPLSKLELOVDSVTSAQOYLITGSEYEL
CFTIPEBRKGSERNLAHCGTRKVTICGQIRPAGITFELNQGKLLDMQDLAIGTDHKVKA
"
complement(7014..7484)
/gene="VC2273"
complement(7014..7484)
/gene="VC2273"
/note="similar to GB:M26839 SP:P04381 GB:X00681 GB:X64395
PID:146976; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="N utilization substance protein B"
/protein_id="AAF95417.1"
/db_xref="GI:9656835"
/translation="WGASVKKPAARRNAFOALQAIYSQITKENVATIEOFLTSGKY
DEEERRAEPALAAIEDVSYFRDLACGVVNHHELSKLPFVSRMODIMNELAL
LRLAYEMTRREDVYKVINAEATELAKVFAEDSKFVNVGLDKAAPVHKKA"
complement(7484..8005)
/gene="VC2274"
complement(7484..8005)
/gene="VC2274"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="phosphatidylylglycerophosphatase A"
/protein_id="AAF95418.1"
/db_xref="GI:9656836"
/translation="MSDPRSNLSLSNPWHLATGFGSGSPIDVGTMTLASVFFYLT
LAQLEWIIAYVAVVAFSLIGIKCOVTSDDQVNDHGSIVDEEAGFWITWLYPLQX
LPVPMKMLVGFVLEFRFMDLKPWIPMLDKRVHGGGIMIDDLVAGVMAISLAI
GYWLEMLV"
complement(5949..6953)
/gene="VC2275"
complement(5949..6953)
/gene="VC2275"
/note="similar to GB:U00096 SP:P77785 PID:1773101
PID:1786619 PID:2329840; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="thiamin-monophosphate kinase"
/protein_id="AAF95419.1"
/db_xref="GI:9656837"
/translation="MQRNQRVMMWFGEFNILIDKYSNRQAQRKDVHLAAGDCAIVK
VPENSRVASIDTVAAGTHFLAQNPAWVAHKALASNTSLAAGATPAWVSLTLTP
EIDDAWLPFGCAFEFLAKYVNVOLIGDTRKGPSITLVYGFLEPQEARLRSAGVLA
GDMVITGDGDSQGLDVILIDPEKRHLIPFADIDLEORHYSITPVIQALVHLASSA
IDISGLAADIOLHILRRSNVGCASIDVSLPLSKLELOVDSVTSAQOYLITGSEYEL
CFTIPEBRKGSERNLAHCGTRKVTICGQIRPAGITFELNQGKLLDMQDLAIGTDHKVKA
"
complement(7014..7484)
/gene="VC2276"
complement(7014..7484)
/gene="VC2276"
/note="similar to GB:M26839 SP:P04381 GB:X00681 GB:X64395
PID:146976; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="N utilization substance protein B"
/protein_id="AAF95420.1"
/db_xref="GI:9656838"
/translation="WGASVKKPAARRNAFOALQAIYSQITKENVATIEOFLTSGKY
DEEERRAEPALAAIEDVSYFRDLACGVVNHHELSKLPFVSRMODIMNELAL
LRLAYEMTRREDVYKVINAEATELAKVFAEDSKFVNVGLDKAAPVHKKA"
complement(7484..8005)
/gene="VC2277"
complement(7484..8005)
/gene="VC2277"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="phosphatidylylglycerophosphatase A"
/protein_id="AAF95421.1"
/db_xref="GI:9656839"
/translation="MSDPRSNLSLSNPWHLATGFGSGSPIDVGTMTLASVFFYLT
LAQLEWIIAYVAVVAFSLIGIKCOVTSDDQVNDHGSIVDEEAGFWITWLYPLQX
LPVPMKMLVGFVLEFRFMDLKPWIPMLDKRVHGGGIMIDDLVAGVMAISLAI
GYWLEMLV"
complement(594

/note="similar to GB:U11391 PID:496182 SP:P51963;
 identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="riboflavin synthase, beta subunit"
 /protein_id="AAF95412.1"
 /db_xref="GI:9656830"
 /translation="MESGSDSHLNRRKGLKMYIEGCFAPNKKIVISRFNSFIN
 ESLSLASIDTLKRNGQISDNNITVVRPGAVEPLVQARAKTDYDAIVSLGVINGK
 GTPFDYVCESEMNKGIAOVSLFESIPAFGVLTVDITDIOAIERAGTRKGAEMALS
 ALEMINVLSLEDS"
 complement(8219 . .9328)
 /gene="VC2269"
 complement(8219 . .9328)
 /gene="VC2269"
 /note="similar to GB:M90094 SP:Q02008 PID:150708;
 identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="3,4-dihydroxy-2-butanone 4-phosphate
 synthase/GTP cyclohydrolase II"
 /protein_id="AAF95413.1"
 /db_xref="GI:9656831"
 /translation="MPTSPQEIIEDIRCKMYTLMDDPDRNEGDILMAEHITPA
 INFATHGRIICLTITKRCRRGLGPMQDNNAOYTTFVYSIEAEGVYTGISA

Query Match 5.4%; Score 68.4; DB 1; Length 11695;
 Best Local Similarity 52.3%; Pred. No. 0.0017;
 Matches 147; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

OY	727	TCAATTTATGCATTATGGCATTAACCTTCCAAATATGTTGACAAAGTCAACAGGTGTTAA	786
Db	11104	TCAATGATGTCCTCGCGGATGAACCTTGCCTGGCGCGGCTTTACCACTTCCTCTAA	11045
OY	787	TCCACCCGTTGGCGCTGTTAGTATATGAAAGTAAAGTATGTTGATGTGGACACTT	846
Db	11044	TCTTATATCGGCTGCGGATCACTCGCGGTGAGCAGATTAGAGGGAAGTTTCCATTT	10985
OY	847	GAGAAAGTGACAAAGCATGCGGAGGTTCAAGCACTTATGCGCAACANAATGCTGA	906
Db	10984	TGCGCGGGGCGCAACCCATGCTGAAGTGCATGCTATGCCCAAGCGGTGAGCTTACG	10925
OY	907	AGGTGCGACGATTTATATACGTTAGAGCCATGTATGTCATTTGGTTCAACACCACTG	966
Db	10924	CGGTGCGACCCCTTATGCACTTAAAGACCTTGTCTCATTTATGTCGACACCGCCCTTG	10865
OY	967	TGTTACAAATTTATGATTTAGATGATGCAANAGTACTAT	1007
Db	10864	TGCTGAAGGCTTATTAAAGCCGGGGTTCGCAAGATGATTT	10824

RESULT 14
 ECNUB 2949 bp DNA BCT 26-JUL-1993
 LOCUS
 DEFINITION E.coli genes nusB (ssyB) and 3 orfs.
 ACCESSION X64395.547077
 VERSION X64395.1 GI:42147
 KEYWORDS nusB (ssyB) gene.
 SOURCE Escherichia coli.
 ORGANISM Escherichia coli.
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.
 1 (bases 1 to 2949)
 Ito, K.
 Direct Submission
 Submitted (06-FEB-1992) K. Ito, Inst for Virus Research, Kyoto
 University, Sakyo-Ku, Kyoto 606-01, JAPAN
 2 (bases 1 to 2949)
 Taura, T., Ueguchi, C., Shiba, K. and Ito, K.
 Insertional disruption of the nusB (ssyB) gene leads to
 cold-sensitive growth of Escherichia coli and suppression of the
 secY24 mutation
 Mol. Gen. Genet. 234 (3), 429-432 (1992)

FEATURES	source	location/qualifiers
MEDLINE 930204316		1..2949
		/organism="Escherichia coli"
		/strain="K12"
		/db_xref="taxon:562"
		/map="9.5 min"
		330..357
misc_feature		/note="promoter-like sequence p1"
		333..362
misc_feature		/note="promoter-like sequence p1"
		391..840
CDS		/note="orf1"
		/codon_start=1
		/transl_table=1
		/protein_id="CAA45734.1"
		/db_xref="GI:42148"
		/db_xref="GI:42148"
		/db_xref="SWISS-PROT:P25538"
		/translation="MHCPCFADVTKIDSLTVGEGSSVRRRCIOVNEFTFFVFAV ELVWPRVKSNDVNEPNEEKLRSGMLRALEKRRVSSDDVEMALNHKLSQRAIGED VPSMIGNLVMEQLKDKKAVIRFASVYRSFEDIKEGEIARLED"
		844..1947
CDS		/note="orf2, homologue to B.subtilis ribS"
		/codon_start=1
		/transl_table=1
		/protein_id="CAA45735.1"
		/db_xref="GI:581147"
		/db_xref="SWISS-PROT:P25539"
		/translation="MODEEYMARALKLAQRGRFTTHPNVNCVIVKDEIVGEYHQ RAGPHAEVHALRMOAGKAKATAYVLEPCSHHGRRPPCCDALIAGVAVASMODD RPNVAGRLVYRLQOAGIDVHGLMSAEQNLKGLKRMPTGEPYIOIKASLDR TNAAGSQRITTSFOARDVOLRAQSHALITTSATYLADDPALTVKRSLEDETOAL YPODNLROPRIIVDSQNRVTPVRIYQDPETVFAVQEDSRWEMPTVRLIPEHK GHLDVLVLMQGLKQOINSIWENAGPTLAGLDGLVDELIYIAPRLGSDARGLC TLPEKLEADAPQRFKEIRHVGDVCLHIVGA"
		1957..1986
misc_feature		/note="promoter-like sequence p2"
		2036..2506
CDS		/note="orf3, homologue to B.subtilis ribH"
		/codon_start=1
		/transl_table=1
		/protein_id="CAA45736.1"
		/db_xref="GI:42150"
		/db_xref="SWISS-PROT:P25540"
		/translation="MNIIEANVATPDARVAITARFNNFINDSLLEGALDKRIGOV KDEIITVWVGALVELLAGALAKTKYPAVIALGVIRGTHAFEYVAGASNGILA HVADDSIEPAVAGVLTTESIEQAIERAGTAKNGAGAAITALEMINVLAKRA"
		2526..2945
gene		/gene="nusb (ssyb)"
		2526..2945
CDS		/gene="nusb (ssyb)"
		/codon_start=1
		/transl_table=1
		/protein_id="CAA45737.1"
		/db_xref="GI:581148"
		/db_xref="SWISS-PROT:P04381"
		/translation="MKPARRRARCECAVQALYSWQSLNDIADVEYOFLEADYKVDV VLEFRELIGAVTNTAYLADGIMKPYLSRLLEEIGQVSRKAVLRALRYELSRSDVPYKVV AINRAIFLAKSFGEASHDKFVNGVLDAAPVIRNRK"
		2945..3245
BASE COUNT		708 a 739 c 825 g 677 t
ORIGIN		
Query Match	5.3%:	Score 67.6; DB 2; Length 2949;
Best local Similarity	52.3%:	Pred. No. 0.003;
Matches 145:	Conservative 0;	Mismatches 132; Indels 0; Gaps 0;
QY	729	AATTATGATTATGCGATTCAACTCCAAATATGTCACAAAGTCACAGAGTCGTTAATC 788
DB	857	ATTACATGGCGCGGGCGCTTAAGCTGGCGGCAACGACGAGCTTTTACCACGATCCCAACC 916
XY	789	CACCCGTTGGCGCGTGTGTACTTAATGAGGTAGGATGTTGGTATTGGTCACACTTGA 848

REFERENCE	TITLE	AUTHORS	JOURNAL
917	CGAATGTCGGGTGCGTCAATTGTCAAAAGATGGCGAAATGTGCGGGAAGTTACCAACCAAC		976
849	GAAGAAGTGACACAGCATGCGGAGSTTCAAGACTTGATATGCGACACACANAAATGCTGAG		908
977	GTGGCGGGTGAACCCATGCTCCCAAGATACACGGTGGCGGATGAGGGGTTGAAAAACCCAAAG		1036
909	GTGGACAGATTTATATTAACGTATAGAGCCATGAGCATTTTGGTTCAACACCACTGTG		968
1037	GTGGACACCGCTATGTCACATCGAACCCCTTACCATCATGATGTCATACGCCACACGTCCT		1096
969	TTACAAATATTTATTTGATTGTAGATAGCANNAAGT		1005
1097	GTGACGCACTCATCGCGCTGCGGAGCGCGCTGTGT		1133
RESULT 15			
AE000148			
LOCUS	AE000148	11692 bp	DNA
DEFINITION	Escherichia coli K-12 MG1655 section 38 of 400 of the complete genome.		BCT 12-NOV-1998
ACCESSION	AE000148	U00096	
VERSION	AE000148.1	GI:1786614	
KEYWORDS			
SOURCE	Escherichia coli.		
ORGANISM	Escherichia coli.		
REFERENCE	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.		
AUTHORS	1 (bases 1 to 11692)		
JOURNAL	Blattner, F.R., Plunkett, G. III, Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Colodner-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B. and Shao, Y.		
MEDLINE	The complete genome sequence of Escherichia coli K-12		
REFERENCE	Science 277 (5331), 1453-1474 (1997)		
AUTHORS	2 (bases 1 to 11692)		
JOURNAL	Blattner, F.R.		
TITLE	Direct Submission		
REFERENCE	Submitted (16-JAN-1997) Guy Plunkett III, Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.		
AUTHORS	Email: ecoligenetics.wisc.edu Phone: 608-262-2534 Fax: 608-263-7459		
JOURNAL	3 (bases 1 to 11692)		
REFERENCE	Blattner, F.R.		
AUTHORS	Direct Submission		
TITLE	Submitted (02-SEP-1997) Guy Plunkett III, Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.		
JOURNAL	Email: ecoligenetics.wisc.edu Phone: 608-262-2534 Fax: 608-263-7459		
REFERENCE	4 (bases 1 to 11692)		
AUTHORS	Plunkett, G. III.		
JOURNAL	Submitted (13-OCT-1998) Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA		
COMMENT	This sequence was determined by the E. coli Genome Project at the University of Wisconsin-Madison (Frederick R. Blattner, director). Supported by NIH grants HG00301 and HG01428 (from the Human Genome Project and NCHGR). The entire sequence was independently determined from E. coli K-12 strain MG1655. Predicted open reading frames were determined using Genemark software, kindly supplied by Mark Borodovsky, Georgia Institute of Technology, Atlanta, GA, 30332 [e-mail: mark@amber.gatech.edu]. Open reading frames that have been correlated with genetic loci are being annotated with CG Site Nos., unique ID nos. for the genes in the E. coli Genetic Stock Center (CGSC) database at Yale University, kindly supplied by Mary Berlyn. A public version of the database is accessible (http://cgsc.biology.yale.edu). Annotation of the genome is an ongoing task whose goal is to make the genome sequence more useful by correlating it with other data. Comments to the authors are appreciated. Updated information will be available at the E. coli Genome Project's World Wide Web site (http://www.genetics.wisc.edu). ***The E. coli K-12 sequence and its annotations are periodically updated; this is version M54. No		

sequence changes. Annotation updates: updated gene identifications and products: all new functional assignments courtesy of Monica Riley; added promoters, protein binding sites, and repeated sequences described in reference 1. The unique numeric identifiers beginning with a lowercase 'b' assigned to each gene (protein- or RNA-encoding) are now designated as gene synonyms instead of labels. This should allow them to be searched for in Entrez as gene names.

FEATURES

```
source
1. .11692
/organism="Escherichia coli"
/strain="K-12"
/sub_strain="MG1655"
/db_xref="taxon:562"

promoter
17. .46
/note="factor Sigma70; predicted +1 start at 432203"
complement(38. .67)
/note="factor Sigma70; predicted +1 start at 432181"

gene
76. .525
/gene="ybad"
/note="b0413"
76. .525
/gene="ybad"

CDS
76. .525
/function="orf; Unknown"
/note="o149; 100 pct identical to yBAD_ECOLI SW: P25538"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAC73516.1"
/db_xref="GI:1786615"
/translation="MHCPCFCAATDVIVDSRLVGGSSVRRRCQCLVCNERETFEVA
ELVMPKVSNDVREPPNEKLSGMLRLKRPVSDVDYEMAINIKSOLRATGERE
VBSKMGNLVMEGLKRLDVAVIRFASVRSFEDIDFEDELRLD"
complement(88. .116)
/note="factor Sigma70; predicted +1 start at 432231"
418. .446
/gene="ypad"
/note="factor Sigma32; predicted +1 start at 432603"
529. .1632
/gene="ribD"
/note="b0414"
529. .1632
/gene="ribD"

CDS
529. .1632
/function="enzyme; Biosynthesis of cofactors, carriers:
Riboflavin"
/note="o367; 100 pct identical to RIBG_ECOLI SW: P25539;
alternate gene name ribG"
/codon_start=1
/transl_table=11
/product="bifunctional pyrimidine deaminase/reductase in
pathway of riboflavin synthesis"
/protein_id="AAC73517.1"
/db_xref="GI:1786616"
/translation="MODEVYMARALKIAGRGRTTPNPNVGVICDGEIVSEGYHO
RAGEPRAEYHALPMAGEKAKATAYTTLEPCSHGRTPECCDALLIAGARVAYASOD
PNPQVAGRLYRLQAGIDVSHGLMSEALQKRGFLKRRRTGFPYIQKLKGLSDGR
TMAAGSEOMITSPQARDVOLRAOSHAILTSASAVLADDPALIVRWSELDEOTAL
PQOONRQPIRIIVDSNRVTPVHRIVOQGFEMFARLDEDSREMPETVATLLIPKH
GHLDIVVLMMOGKQOINSIWEPAGPTLIGALLQDLVDLLIYIAPKLLGSDARLCL
TLPGLKLDADPQFKETIRHVPDVCMLHVCA"
1343. .1571
/gene="ribD"
/note="factor Sigma32; predicted +1 start at 433728"
1721. .2191
/gene="ribH"
/note="b0415"
1721. .2191
/gene="ribH"
/EC_number="2.5.1.9"
/function="enzyme; Biosynthesis of cofactors, carriers:
Riboflavin"
```

```
promoter
2211. .2630
/note="factor Sigma54; predicted +1 start at 434290"
/gene="nusb"
/note="b0416"
2211. .2630
/gene="nusb"

CDS
2211. .2630
/function="factor; RNA synthesis, modification, DNA
transcription"
/note="o139; 100 pct identical to NUSB_ECOLI SW: P04381;
alternate gene name ssyB"
/codon_start=1
/transl_table=11
/product="transcription termination; L factor"
/protein_id="AAC73519.1"
/db_xref="GI:1786618"
/db_xref="GI:1786618"
/translation="MKPARRRARCCAVQALYSQMDIADVEYQFLAEQDVKDV
VYFRELQAGVANTFAYLDGMLKPYLSRLEEGOVERKVRIRALVYELSKRSQVYK
ATNPELEIASKSGAEDSHKFFVNGVLDKAPVIRPKK"
2642. .2668
/note="factor Sigma70; predicted +1 start at 434825"
2660. .2687
/note="factor Sigma70; predicted +1 start at 434844"
2679. .2706
/note="factor Sigma70; predicted +1 start at 434863"
2708. .3685
/gene="thiL"
/note="b0417"
2708. .3685
/gene="thiL"

CDS
2708. .3685
/function="enzyme; Biosynthesis of cofactors, carriers:
thiamin"
/note="o325; 56 pct identical to 273 aa of GB:
HIU32756_3ACCESSION: U32756 (345 aa)"
/codon_start=1
/transl_table=11
/product="thiamin-monophosphate kinase"
/protein_id="AAC73520.1"
/db_xref="GI:1786619"
/db_xref="GI:1786619"
/translation="MACGEFSLIARYFDRVRSSRLDVELGIDDCALNIPKQTLAI
SDTLVAGNHFLPDIDPADLAYKALAVNSDLAAGADPAMLTALLTPDVEAMLES
FSDSLFDLNNYDMOIGDGTTRGPLSMIGIHGVPQMRALTRSGAKGDMIVYTG
PBDASAAGLATIONRLOVADKADADYLKHHLPSPRIIOGLANLANSADLISGLI
SOLGHIVKASDGCARIDALLPSPALSRLVHPEDRALRGSGEDYELCFVPELNR
GLADVAGLHVSPTFCIGMTADIGLCTIRGEPTVLDWKGYDHPAT"
3663. .4181
/gene="pgpA"
/note="b0418"
3663. .4181
/gene="pgpA"

CDS
3663. .4181
/gene="pgpA"
/function="enzyme; Macromolecule synthesis, modification:
phospholipids"
/note="o172; 50 pct identical (2 gaps) to 155 residues of
hypothetical protein YD06_HAEIN SW: P44157 (163 aa)"
/codon_start=1
/transl_table=11
/product="phosphatidylglycerophosphatase"

Query Match
5.3%; Score 67.6; DB 1; Length 11692;
```

Best Local Similarity 52.3%; Pred. No. 0.0024;
Matches 145; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

```

OY 729 AATTATGATATATCGATTCACCTTCCAAATATGTGACAGGTCANACAGGTGTTATC 788
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 542 ATTACATGCGCGCGCGCTAAAGCTGGCGCAACGAGAGCGTTTACACGCAATCCCAACC 601
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 789 CACCGGTGGCGCGCTGTGTAGTTAATGAAGTAGATGTTGTATTTGTCACACTTGA 848
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 602 CGAATGTCGGGTGCGTCATTGTCAAGATGCGAAATTTGCGTGAGGTTACCAACCAAC 661
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 849 GAAAAGGTGACAAGCATGGGAGGTTCAAGCACTTGATATGGCACAAACANAAATGCTGAAG 908
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 662 GTGCGGTGAACCAACATGCGCAAGTACACGCGTGGGTATGGCGGGTGAAGCAAGCAAG 721
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 909 GTGCGACGATTTATTTAGCTTAGAGCCATGTATGTCATTTGTTCAACACCAACCTGTG 968
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 722 GTGCGACGCGCTATGTACACTGAAACCTGTAGCCATCATGTGTCAGCCACCGTGTGT 781
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 969 TTAACAAATTTATTTAGTAGATAGCANAAAGTAGT 1005
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 782 GTGACGCACTCATCGCGCGCTGGCGTAGCGCGGTGTGT 818
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: November 19, 2000, 04:26:04
Job time: 12947 sec

PI Lonetto MA, Nicholas RO, Palmer LM, Pratt JM, Reichard RW;
PI Rosenberg M, Triani CM, Ward JM, Warren RL;
DR WPL. 1999-347572/29.
DR P-PSDB: Y15914.
PT New *Staphylococcus aureus* polypeptide and polynucleotide useful in
PT the treatment of gastric ulcer and gastritis
PS Claim 24; Page 5-6; 48pp: English.
XX
XX The present sequence encodes a *Staphylococcus aureus* pyrimidine
CC deaminase and pyrimidine reductase (ribG) polypeptide. *Staphylococcus*
CC *aureus* ribG and its antagonists are used to treat individuals in
CC need of them. Disease related to expression or activity of ribG can
CC be determined by analysing the nucleic acid sequence encoding ribG
CC or detecting the ribG polypeptide in a sample. ribG can also be used
CC to identify antagonists or agonists. ribG, or its related nucleic acid,
CC also has use as a vaccine to induce an immunological response in an
CC animal. Antimicrobial compounds (e.g. agonists and antagonists of ribG),
CC especially broad-spectrum antibiotics, may be of use in the treatment
CC of *Helicobacter pylori* infection. This should decrease the advent of
CC *H. pylori*-induced cancers, such as gastrointestinal carcinoma. The
CC treatment should also cure gastric ulcers and gastritis.
SQ Sequence 1269 BP: 435 A; 188 C; 256 G; 371 T; 19 other;

Query Match 98.5%; Score 1250; DB 20; Length 1269;
Best Local Similarity 100.0%; Pred. No. 6,7e-291;
Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AANCAACCAATCCNATTGGAGAGNAATCCAAATCAATCCCGANCCCAATCCAGTTAA 60
DB 1 aancaccaatccnatggagagaaacccaatcaatcccganccccaatccaaat 60
OY 61 TTAAGTCCAAAGTTTGGAACTTACCAATATGATTCGATGAGGTCAATGNCANCG 120
DB 61 ttaagtccaaagtttggaaacttaccatcaatgatccgatgaggtcaaatgncanng 120
OY 121 GTGTTAATAACTAGCAAAATGNTGNAATAGATAGTAGANCAAGTTGGCATACAGTNT 180
DB 121 gttgtaataactagcaaaatgntgnaatagatagtagancaaagtggcatacagtnt 180
OY 181 CTCNATATATATGCTTATGAAATGCAATGACGACAAACAGCGCACTTCAATCAATAT 240
DB 181 ctcnatataatgcttataatgcaatgacgacaaacagcgcacttcaatcaat 240
OY 241 TTGTACTAGAGATATTAAGAAAGCGCTATAGAGAGAAATTAAGTTGATTTTA 300
DB 241 ttgtactagagatattaagaaagcgctatagagagaaatthaagttgatttta 300
OY 301 ATGTCTGTAGTAGAATCATATCAATGAGATGCCATATAGTACTGATTAATTAAT 360
DB 301 atgtctgtatagtagaatacatatcaatgagatgccatatagtactgattataat 360
OY 361 AAAACCGGTATTAATGTTTTTTAGAAAACATATGATTTTAATGTAAGTGA 420
DB 361 aaaacccggtatthaatgttttttagaaaacatatagtatcatttaagtgtgaa 420
OY 421 TACTAGTCTCAAAATATCTTAACATTTATATATTTCTTCCGGGCGAGGTGAA 480
DB 421 tactagctcaaaatatacttaacatttatataatcttccgggcaagggtgaa 480
OY 481 ATTCCCAACCGGACAGTAATAAGCCGTGACCTGCTAATATGTTTCAATTAAGTGG 540
DB 481 attcccaacccggacagtaataagccgtgacctgcttaataatgtttcataatg 540
OY 541 ATCTAGTGAATTTAGACCCGACACTTAAAGTCTGATGGGGAAGAAATGTAATTT 600
DB 541 atctagtgaatttagacccgacactttaaagtctgatggggaagaaatgttaatt 600
OY 601 GCACAAAGATATGATGCTATTTGTAAAAAATGTGTAACAAATAGCTTATTAACATTA 660

DB 601 cgacaaagataatgtagcgtattgttaaaatggtacaataagcttataacgata 660
OY 661 ATTTTTCCTTTCGATCTTAATTCATGATGAGAGATTTTTTGTATTAGAGTGATCA 720
DB 661 atttttcccttctgcatcttaattcaatgtagagattttgtttlaagaggatca 720
OY 721 TTGAGTCAATTAATGATTTATGATTCGATTCACCTCAATATATGTAAGGCAACAG 780
DB 721 ttgagtcatttaagattatgattatgcattccaacttccaatatgtagaagtcana 780
OY 781 TGTTAATCCACCGCTTGGCGCTGTTGATTAATGAGGTAGATGTTGTTGTTGTC 840
DB 781 tgttaatccacccgttggcgctgttgaattaatgaggaagtagattgttgattgtgc 840
OY 841 ACACTTGAGAAAAGTGCACAGCATGCGGAGTTTAAGCAGCTGATATAGCGACANANA 900
DB 841 acacttgagaaaagtgacacagcatcggaaggttcaagcacttgatagcacacana 900
OY 901 TGCTGAGAGTGGAGATTTATATACGTTAGAGCCATGATCAATTTGTTCAACAC 960
DB 901 tgctgagagtgagatttatatacgttagagccatgatacttcttggttcaacc 960
OY 961 ACCCTGTGTTAACAAAATTTATGATTTGTAAGATAGCANAAAGTAGATTACACANAA 1020
DB 961 accctgtgtaacaaaatttatgattgtaagatagcanaaagtagattacnacaana 1020
OY 1021 ACAATTCGTTGACACACATGCGTGATGACGCTTACGGGCTCCAGGTTATTGAGG 1080
DB 1021 acaatttcgttgacacacatgctgtgatgacgcttaccgggctccaggtattga 1080
OY 1081 TTGAATTTGGTTGATGATGAACGGGCATCACATTAATCCAAAGACTTTTAAAGA 1140
DB 1081 ttgaatttggttgatgatgaacgggcatacatatataccaaagactttttaaaga 1140
OY 1141 AAAAGCAAGCAACTTGGCACAAATTAACAGTGAAGTNTCTTGAAGTTAGATGGTA 1200
DB 1141 aaagcaagcaacttggcacaaatthaacagtgaagntcttgaagtttagatggta 1200
OY 1201 AACAAAGCAATTTGATTAATGACAAAGTCAATGATTTACTAACAAAGGTTAAACAGA 1260
DB 1201 aacaaagcaattgataatgacaaagtcaatgatttactaacaaaggttaacaa 1260
OY 1261 TGCTATAG 1269
DB 1261 tgcataag 1269

RESULT 3
V74453/c
ID V74453 standard; DNA: 7588 BP.
XX
AC V74453:
XX
DT 16-MAR-1999 (first entry)
XX
DE *Staphylococcus aureus* contig SEQ ID #142.
XX
KW Computer readable medium; vaccine; *S. aureus* infection; immunodetection;
KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW skin infection; surgical wound infection; scalded skin syndrome;
KW toxic shock syndrome; ds.
OS *Staphylococcus aureus*.
XX
FH Key Location/Qualifiers
FT misc-feature 661..720
FT
FT
FT
FT misc-feature 2461..2520

These bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"

FT /+tag- b "these bases represent a line of missing text in
FT /note- the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"
FT misc-feature 4261..4320
FT /+tag- c
FT /note- "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"
FT misc-feature 6061..6120
FT /+tag- d
FT /note- "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"
PN EP86519-A2.
XX 30-JUL-1997.
XX 07-JAN-1997; 97EP-0100117.
XX 05-JAN-1996; 96US-0009861.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;
XX Rosen CA;
XX WPI; 1997-374922/35.
DR Polynucleotide(s) and proteins derived from *Staphylococcus aureus*
PT stored on computer readable medium and used in the production of
PT anti-S.aureus vaccines
XX
XX Claim 1; Page 740-744; 3271pp; English.
PS This sequence represents one of 5191 *Staphylococcus aureus* DNA sequences
XX of the invention. The DNA sequences are recorded on a computer readable
XX medium, preferably selected from a floppy or hard disk, random access
XX memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
XX the S.aureus DNA sequences allows putative functions to be assigned so
XX that protein-encoding or regulatory regions of commercial, therapeutic or
XX industrial importance can be obtained. Specifically, sequences which are
XX likely to encode antigens have been identified and these polypeptides can
XX be used in a vaccine composition against S.aureus infection. The
XX polypeptides can also be used in a kit for the immunodetection of
XX S.aureus in a sample. S.aureus is implicated in numerous human diseases,
XX including cellulitis, eyelid infections, food poisoning, osteomyelitis,
XX skin and surgical wound infections, scalded skin syndrome, toxic shock
XX syndrome, etc. Organisms transformed with the DNA sequences can be used
XX for recombinant production of the polypeptides. The new DNA sequences
XX (and their fragments) are useful as primers or probes for isolating
XX homologues of any of the S.aureus DNA sequences contained on the
XX computer readable medium.
XX
XX Sequence 7588 BP; 2333 A; 1341 C; 1061 G; 2613 T; 240 other;

Query Match 70.7%; Score 897.6; DB 18; Length 7588;
Best Local Similarity 88.1%; Pred. No. 3,7e-206;
Matches 1058; Conservative 0; Mismatches 122; Indels 21; Gaps 11;

QY 68 CAAGGTTTGGACATTAACCAATATGATCCGATGAGTCGAATGCAACGCGTTAA 127
DB 6512 CCAAGGTTTGGACATTAACCAATATGATCCGAT-AGGTCGAATGCAACGCGTTAA 6454
QY 128 TAAACTAGCAAAATGTGTGNAATGATAGTAGCAAGATTCGCGATCAGTNTCTCAAT 187
DB 6453 TAAACTAGCAAAATGTGTGNAATGATAGTAGCAAGATTCGCGATCAGTNTCTCAAT 6394

QY 188 ATATGATGCTTTAGATCGAATGAGCAACAGCCAGTATCAATCAATATTTGTAC 247
DB 6393 ATATGATGCTTTAGATCGAATGAGCAACAGCCAGTATCAATCAATATTTGTAC 6334
QY 248 TAGAGATTAATAAGAAAGCGCTCTATAGAGCAAGTGAAGTTGATTTAATGCTG 307
DB 6333 TAGAGATTAATAAGAAAGCGCTCTATAGAGCAAGTGAAGTTGATTTAATGCTG 6274
QY 308 TTGATAGAAATCATATCAATGAGATGCTATAGTACTGATTTATTTAATTAACCG 367
DB 6273 TTGATAGAAATCATATCAATGAGATGCTATAGTACTGATTTATTTAATTAACCG 6214
QY 368 TCATTAATTTGTTTTAGAAAACATATAGTATCATTTTAAATGATGATCTACTAG 427
DB 6213 TCATTAATTTGTTTTAGAAAACATATAGTATCATTTTAAATGATGATCTACTAG 6154
QY 428 TACTCAATATATCTATATACAAATTTATATATAATTTCTTGGGGCAGGGTGAATTCGA 487
DB 6153 TACTCAATATATCTATATACAAATTTATATATAATTTCTTGGGGCAGGGTGAATTCGA 6094
QY 488 ACCGGCAGTAATAAAGCCCTGCAGCTCTAATATGTTTCATTTAGTGGCTGATCTAGT 547
DB 6093 NNN 6034
QY 548 GAGATTTCTAGAGCCGACAGTAAAGTCTGGATGGAGAAAGAAATGTTATTCGACAAA 607
DB 6033 GAGATTTCTAGAGCCGACAGTAAAGTCTGGATGGAGAAAGAAATGTTATTCGACAAA 5974
QY 608 GATTAATGTAGCGTATTTGTAATAATGTGTAACAAATAGGCTTATTTAGCATTAATTTTC 667
DB 5973 GATTAATGTAGCGTATTTGTAATAATGTGTAACAAATAGGCTTATTTAGCATTAATTTTC 5914
QY 668 TCCTTTGATCTTAATTTCAATGATGATGATGATTTTGTGTTATAGAGTATCTTTGAGT 727
DB 5913 TCC-TTGCATCTTAATTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 5856
QY 728 CAATTTATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 787
DB 5855 CAATTTATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5796
QY 788 CCACCGTTGGCGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 847
DB 5795 CCACCGTTGGCGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5736
QY 848 AGAAAGGTGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 907
DB 5735 AGAAAGGTGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5676
QY 908 GGTGCGACGATTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 967
DB 5675 GGTGCGACGATTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5616
QY 968 GTTACAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1027
DB 5615 GTTACAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5556
QY 1028 CGTTAGACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1087
DB 5557 CGTTAGACACAT-GGTAGAGAGCTTACGGGCTCACGGGATGATGATGATGATGATGAT 5499
QY 1088 GCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1147
DB 5498 -----GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5448
QY 1148 AAGCACTTGCACAAATTTACAGTGAAGTNTCTGAAAGTTTACAGTGGTAAACAAAG 1207
DB 5447 AAGCACT--GCCACAAATTTACAGTGAAGTATC--TGCAAGTTTACAGTGGTAAACAAAG 5391
QY 1208 CGAATGATATAGCAAAAGTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1267
DB 5390 ---AATGATATAGCAAAAGTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 5334
QY 1268 A 1268

Db 5333 A 5333

RESULT 4
X59909
ID X59909 standard; DNA; 1029 BP.
XX
XX X59909;
DT 04-AUG-1999 (first entry)
XX
XX Pyrimidine deaminase and pyrimidine reductase (ribg) DNA.
DE
XX Pyrimidine deaminase; pyrimidine reductase; ribg; antagonist;
KM agonist; antimicrobial; antibiotic; Helicobacter pylori infection;
KW H. pylori-induced cancer; gastrointestinal carcinoma; gastric ulcer;
KM gastritis; ss.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT 1..1029
FT CDS /*tag- a
FT /product= "ribg"
FT
FT
PN MO926475-A1.
XX
PD 03-JUN-1999.
XX
PF 23-NOV-1998; 98MO-US24857.
XX
PR 25-NOV-1997; 97US-0978456.
XX
XX (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI Black MT, Burnham MKR, Fedon JC, Hodgson JE, Knowles DJC;
PI Lonetto MA, Nicholas RO, Palmer LM, Pratt JM, Reichard RW;
PI Rosenberg M, Trani CM, Ward JM, Warren RL;
DR MPI: 1999-347572/29.
DR P-PSDB: Y15913.
XX
XX
PT New Staphylococcus aureus polypeptide and polynucleotide useful in
the treatment of gastric ulcer and gastritis
XX
PS Claim 6; Page 5; 48pp; English.
XX
XX The present sequence encodes a Staphylococcus aureus pyrimidine
deaminase and pyrimidine reductase (ribg) polypeptide. Staphylococcus
aureus ribg and its antagonists are used to treat individuals in
need of them. Disease related to expression or activity of ribg can
be determined by analysing the nucleic acid sequence encoding ribg
or detecting the ribg polypeptide in a sample. ribg can also be used
to identify antagonists or agonists. ribg, or its related nucleic acid,
also has use as a vaccine to induce an immunological response in an
animal. Antimicrobial compounds (e.g. agonists and antagonists of ribg),
especially broad-spectrum antibiotics, may be of use in the treatment
of Helicobacter pylori infection. This should decrease the advent of
H. pylori-induced cancers, such as gastrointestinal carcinoma. The
treatment should also cure gastric ulcers and gastritis.
XX
XX
SQ Sequence 1029 BP; 376 A; 162 C; 198 G; 293 T; 0 other;

Query Match 27.8%; Score 352.2; DB 20; Length 1029;
Best Local Similarity 88.4%; Pred. NO. 1.4e-75;
Matches 473; Conservative 0; Mismatches 44; Indels 18; Gaps 8;

QY 734 ATGATTATCGATTCACTTCCAAATATGATACAGAGTGTATATCCACC 793
|||||
Db 1 atgattatcgattcaacttgcgaataatgtaaacaggtgttaaccaccc 60

QY 794 GTTGCCCTGTCTAGTTAATGAAGTAGATTGTGATGTGCACCTTGAGAAA 833
|||||
Db 61 gtggcgctgttgaagtaagagtagattgtgtatgtgtgacacttgagaaaa 120
QY 854 GGTGACAGCATGGCGAGGTTCAAGCATTGATATGGCACACANNAATGCTGAAGTGGC 913
|||||
Db 121 ggtgacaagcatgcggaggttcaagcacttgatatgacacaaaatgtctgaagtcg 180
QY 914 ACGATTATATTCAGTATGAGCCATGATGATTTTGGTTCAACACACCTGGTTAAC 973
|||||
Db 181 acgattatatacgttaagaccatgtagtcatcttggttcaacacacccgtgtaac 240
QY 974 AAATTATGATTGTATGATAGCANAAGTAGATTATACNACANNAAGCAATTCGGTTAG 1033
|||||
Db 241 aaaatttattatgtaagaagaacaaaagtgtat-taagcaacaaagacatt-cgttag 298
QY 1034 ACACACATGGGTGATGAGACGTTACGGGCTCCACGGTATTGAGGTTAAATTGGCTTG 1093
|||||
Db 299 acaacacat-gtgtatgagacttaacggtcaccgtaattgagttgattgctt----- 351
QY 1094 GATGATGAAGGGCATACAAATATATACCAAGACTTTTAAAGCAAAAGCAAGCAA 1153
|||||
Db 352 gatgatgaacgggcatcacattataccaaagacttlt--aaagcaaaagaaagcaa 408
QY 1154 CTTCGCCAATAATTACAGTAAAGTNTCTGAAAGTTAGATGGGTAAACAAAGCAATT 1213
|| | |||||
Db 409 ct--gccacaattacagtgtaaaagtatc-tgcaagtttaagtgttaacaaagc---aat 462
QY 1214 GATATGACAAAGTCAATGATTTACTAACAAAGGTTAAACAGATGCTATA 1268
|||||
Db 463 gataatgacaaagtcaatgattactaacaaagagttaaacaagtgctata 517

RESULT 5
X59913
ID X59913 standard; DNA; 3336 BP.
XX
XX
AC X59913;
XX
DT 04-AUG-1999 (first entry)
XX
XX Riboflavin biosynthesis operon containing the ribg gene.
DE
XX Pyrimidine deaminase; pyrimidine reductase; ribg; antagonist;
KM agonist; antimicrobial; antibiotic; Helicobacter pylori infection;
KW H. pylori-induced cancer; gastrointestinal carcinoma; gastric ulcer;
KM gastritis; ss.
XX
XX Staphylococcus aureus.
OS
XX
XX MO926475-A1.
XX
XX
XX 03-JUN-1999.
PD
XX
XX 23-NOV-1998; 98MO-US24857.
PF
XX
XX 25-NOV-1997; 97US-0978456.
PR
XX
XX (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI Black MT, Burnham MKR, Fedon JC, Hodgson JE, Knowles DJC;
PI Lonetto MA, Nicholas RO, Palmer LM, Pratt JM, Reichard RW;
PI Rosenberg M, Trani CM, Ward JM, Warren RL;
DR MPI: 1999-347572/29.
XX
XX
PT New Staphylococcus aureus polypeptide and polynucleotide useful in
the treatment of gastric ulcer and gastritis
XX
PS Example 2; Page 30-31; 48pp; English.
XX
XX The present sequence represents the Staphylococcus aureus
CC riboflavin biosynthesis operon containing the pyrimidine

CC deaminase and pyrimidine reductase (ribG) gene. *Staphylococcus*
 CC aureus ribG and its antagonists are used to treat individuals in
 CC need of them. Disease related to expression or activity of ribG can
 CC be determined by analysing the nucleic acid sequence encoding ribG.
 CC or detecting the ribG polypeptide in a sample. ribG can also be used
 CC to identify antagonists or agonists. ribG, or its related nucleic acid,
 CC also has use as a vaccine to induce an immunological response in an
 CC animal. Antimicrobial compounds (e.g. agonists and antagonists of ribG),
 CC especially broad-spectrum antibiotics, may be of use in the treatment
 CC of *Helicobacter pylori* infection. This should decrease the advent of
 CC H. pylori-induced cancers, such as gastrointestinal carcinoma. The
 CC treatment should also cure gastric ulcers and gastritis.

CC Sequence 3336 BP; 1202 A; 460 C; 677 G; 997 T; 0 other;

Query Match: 27.8%; Score 352.2; DB 20; Length 3336;
 Best Local Similarity 88.4%; Pred. No. 1.9e-75;
 Matches 473; Conservative 0; Mismatches 44; Indels 18; Gaps 8;

```

OY 734 ATGATTTATGCGATTCCAAATATGATGACAAAGTCANACAGGCTTATCCAGCC 793
DB 1 atgattatgcatcattcaacttgcataatgacaaagtcataaccacc 60
OY 794 GTTGGCGCTGTGTAGTAAATGAGGTAGATGTTGTTGTTGTCACACTTGAGAAA 853
DB 61 gtggcgctgtgttagtaagtagatgttggtagtggtagacacttgagaaa 120
OY 854 GGTGACAAAGCATGCGAGGTTCAAGCATTGATGACACANANATCTGAAGTGG 913
DB 121 ggtgacaacatgcygaggttcaagcattgatalgcaacaataatgtaagttcg 180
OY 914 ACGATTTATATACGTTAGAGCCATGATGATTTGTTGTTCAACACCCCTGTGTAC 973
DB 181 acgattatatacgttagaagccatgtagtcatltagtcaaccaccctgtgtaac 240
OY 974 AAAATTTATGATTGATAGTACAGCANAAGTAGTATTACNACANAGACATTCGTTAG 1033
DB 241 aaattattgattgtaagtagaagaagtagta-tacgacaacaagaacatt-cgttag 298
OY 1033 ACACACATGGGTGATGAGACGTTACGGGGCTCCACGCTATTGAGGTTGATTCGTTG 1093
DB 299 acaccacat-ggtgatgagacgttcaagcgtcacgtagtgaagttgaagcgtt----- 351
OY 1094 GATGATGAACGGGATCAATATTTACCAAGACTTTTAAAGCAAAAGCAAGCAA 1153
DB 352 gatgatgaacgggacatcaaatatataccaagactttt---aaagcaaaagcaagcaa 408
OY 1154 CTGGCCACAAAATTCAGTGAAGTNTCTTGAAGTTTATGATGGGTAAACAAAGCAATT 1213
DB 409 ct-gccacaatttacagtgaaagatc-tgcaagtttagatgtaacaagcg---aat 462
OY 1214 GATTAATGCAAAAGTCAATGATTTACTATCAAAAGAGGTTAAACAGATCTATA 1268
DB 463 gataatgcaaaagtcataatgatactactaacaagaaggttaacaagaatgctata 517

RESULT 6
X77007
ID X77007 standard; DNA; 3336 BP.
XX AC X77007;
XX DT 06-AUG-1999 (first entry)
XX DE S. aureus riboflavin biosynthesis operon.
XX XX
XX KM GTP cyclohydrolase II; ribA; diagnosis; vaccine; bacterial adhesion;
XX KM Helicobacter pylori infection; gastrointestinal carcinoma; gastric ulcer;
XX KM gastritis; wound healing; infection; drug screening; therapy;
XX KM riboflavin biosynthesis operon; ss.
XX OS Staphylococcus aureus.

```

XX W09926646-A1.
 XX 03-JUN-1999.
 XX 24-NOV-1998; 98WO-US25068.
 XX 25-NOV-1997; 97US-0977554.
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 XX Black MT, Burnham MKR, Debouck C, Fedon JC, Hodgson JE;
 XX Jaworski DD, Knowles DJC, Lonetto MA, Mooney J, Nicholas RO;
 XX Palmer LM, Pratt JM, Reichard RM, Rosenberg M, Traini CM;
 XX Wang M, Ward JM, Warren RL, Zhong Y;
 XX WPI. 1999-357736/30.

PT New *Staphylococcus aureus* ribA polypeptide - and methods for using
 PT the polypeptide to screen for antibacterials

PS Claim 24; Page 36-37; 53pp: English.

CC This sequence represents the *S. aureus* riboflavin biosynthesis operon,
 CC which contains the ribA coding sequence of the invention. ribA belongs to
 CC the GTP cyclohydrolase II family. The polynucleotides and polypeptides
 CC may be employed as research reagents and material for the discovery of
 CC treatments and diagnostics for diseases, particularly human diseases.
 CC They can be used for diagnosis of the disease and staging of disease, and
 CC as reagents in differential screening methods. The polynucleotides may be
 CC used as a source for hybridisation probes, and for screening of genetic
 CC mutations, serotype, and identification, and for organism chromosome
 CC identification. The polypeptides can be used to produce antibodies. The
 CC polypeptides can also be used in vaccine formulations, and to identify
 CC agonists and antagonists. These are used to prevent, inhibit or treat
 CC diseases, particularly of *Helicobacter pylori* infections, such as
 CC gastrointestinal carcinoma. Gastric ulcers and gastritis may also be
 CC treated. The polypeptides can also be used to treat wounds and
 CC in-dwelling devices to prevent bacterial adhesion and infection, and to
 CC block ribA protein-mediated mammalian cell invasion. The frequency of
 CC *Staphylococcus* infections has risen dramatically, and it is no longer
 CC common to find isolated *S. aureus* strains that are resistant to standard
 CC antibiotics. The ribA products of the invention can be used screen for
 CC new antibacterial compounds that may target these resistant bacteria.

XX Sequence 3336 BP; 1202 A; 459 C; 678 G; 997 T; 0 other;

Query Match: 27.8%; Score 352.2; DB 20; Length 3336;
 Best Local Similarity 88.4%; Pred. No. 1.9e-75;
 Matches 473; Conservative 0; Mismatches 44; Indels 18; Gaps 8;

```

OY 734 ATGATTTATGCGATTCCAAATATGATGACAAAGTCANACAGGCTTATCCAGCC 793
DB 1 atgattatgcatcattcaacttgcataatgacaaagtcataaccacc 60
OY 794 GTTGGCGCTGTGTAGTAAATGAGGTAGATGTTGTTGTTGTTCAACACTTGAGAAA 853
DB 61 gtggcgctgtgttagtaagtagatgttggtagtggtagacacttgagaaa 120
OY 854 GGTGACAAAGCATGCGAGGTTCAAGCATTGATGACACANANATCTGAAGTGG 913
DB 121 ggtgacaacatgcygaggttcaagcattgatalgcaacaataatgtaagttcg 180
OY 914 ACGATTTATATACGTTAGAGCCATGATGATTTGTTGTTCAACACCCCTGTGTAC 973
DB 181 acgattatatacgttagaagccatgtagtcatltagtcaaccaccctgtgtaac 240
OY 974 AAAATTTATGATTGATAGTACAGCANAAGTAGTATTACNACANAGACATTCGTTAG 1033
DB 241 aaattattgattgtaagtagaagaagtagta-tacgacaacaagaacatt-cgttag 298
OY 1034 ACACACATGGGTGATGAGACGTTACGGGGCTCCACGCTATTGAGGTTGATTCGTTG 1093

```


DB	463	gataatgacacaagtcacatgattactacaacaagaaggttaacaagaatgctcata	517
RESULT	9		
ID	T83787		
XX	T83787	standard; DNA; 501 BP.	
XX			
AC	T83787;		
XX			
DT	16-JUL-1998	(first entry)	
DE			
XX	DNA encoding a Staphylococcus aureus protein of unknown function.		
XX			
XX	Staphylococcus aureus protein; ribozyme; antisense sequence; control;		
KW	Staphylococcal gene; regulatory element; bacterial gene expression;		
KW	vaccine; Staphylococcal infection; food poisoning; scaled skin syndrome;		
XX	toxic shock syndrome; ss.		
XX			
OS	Staphylococcus aureus.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	167..382	
FT		/*tag= a	
XX			
PN	W09730070-A1.		
XX			
PD	21-AUG-1997.		
XX			
PF	19-FEB-1997;	97WO-US02318.	
XX			
PR	20-FEB-1996;	96US-0011888.	
XX			
PA	(SMIK) SMITHKLINE BEECHAM CORP.		

DR WPI: 1997-424969/39.

XX Novel polypeptide(s) from *Staphylococcus aureus* strain WCUH29 - used
PT to isolate antimicrobial compounds, and in vaccines against *S.*
PT aureus infection
XX
PS Claim 9; Page 673: 989pp; English.
XX
XX The present sequence encodes a *Staphylococcus aureus* protein of
CC unknown function. The present sequence was isolated from a
CC library of clones of *S. aureus* WCUH 29 in *Escherichia coli*. The DNA
CC sequence can be used in the construction of ribozymes and antisense

CC sequence is also useful as a source of regulatory elements for the
CC control of bacterial gene expression. The encoded protein may be used
CC to produce vaccines to enable a host to produce specific antibodies
CC with antibacterial action. These vaccines and antibodies would protect
CC a host against invasion by *S. aureus*, and conditions relating to
CC Staphylococcal infection, e.g. Staphylococcal food poisoning, scaled
CC skin syndrome, and toxic shock syndrome.

xx Sequence 501BP; 189 A; 77 C; 94 G; 133 T; 8 other;
xx

Query Match	18.0%	Score 229;	DB 18;	Length 501;
Best Local Similarity	89.7%	Pred. No. 3	9e-46;	
Matches 261; Conservative	0;	Mismatches 27;	Indels 3;	Gaps

Dd
06 ccaaggttgcgaacttaacaatgatgcatc-aggccaatgcacaaacgctgttaa 270

Db 271 taacttcgaatcgtgtgaaatgtagtagaacaagttgcgcatacagttatcctaatt 330

QY 188 ATATGATGCTTGAATCGAATGAGCAACAGCGCGATTACATCATATTTGTAAC 247
 Db 331 atatgctgcttgaatcgaaatgagcaacacgacgacttaacaataattgtgaac 390
 QY 248 TAGAATATATAAGAGACGCTCTATAGAGCAATTTGAGGTTTG -ATTTTAATGTC 305
 Db 391 taganagataaangaaagcctcagagagaagaattgaagtttgatttncigtct 450
 QY 306 TGTTAGTAGAATCATATCATAGATGCTATAGTACTAGATTAATTA 356
 Db 451 tgttagtaagcattacatcatgagatgctcctatngctcctcagattatatta 501
 RESULT 10
 V75429
 ID V75429 standard: DNA: 223 BP.
 AC V75429;
 XX
 DT 16-MAR-1999 (first entry)
 DE Staphylococcus aureus contig SEQ ID #118.
 XX
 KM Computer readable medium: vaccine; S.aureus infection; immunodetection;
 KM cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
 KM skin infection; surgical wound infection; scalded skin syndrome;
 KM toxic shock syndrome; ds.
 XX
 OS Staphylococcus aureus.
 XX
 PN EP786519-A2.
 XX
 PD 30-JUL-1997.
 XX
 PF 07-JAN-1997; 97EP-0100117.
 XX
 PR 05-JAN-1996; 96US-0009861.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;
 PI Rosen CA;
 XX
 DR WPI; 1997-374922/35.
 XX
 PT Polynucleotide(s) and proteins derived from Staphylococcus aureus
 PT stored on computer readable medium and used in the production of
 PT anti-S.aureus vaccines
 PS
 PS Claim 1; Page 1842; 3271pp; English.
 XX
 CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences
 CC of the invention. The DNA sequences are recorded on a computer readable
 CC medium, preferably selected from a floppy or hard disk, random access
 CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
 CC the S.aureus DNA sequences allows putative functions to be assigned so
 CC that protein-encoding or regulatory regions of commercial, therapeutic or
 CC industrial importance can be obtained. Specifically, sequences which are
 CC likely to encode antigens have been identified and these polypeptides can
 CC be used in a vaccine composition against S.aureus infection. The
 CC polypeptides can also be used in a kit for the immunodetection of
 CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,
 CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
 CC skin and surgical wound infections, scalded skin syndrome, toxic shock
 CC syndrome, etc. Organisms transformed with the DNA sequences can be used
 CC for recombinant production of the polypeptides. The new DNA sequences
 CC (and their fragments) are useful as primers or probes for isolating
 CC homologues of any of the S.aureus DNA sequences contained on the
 CC computer readable medium.
 XX
 SO Sequence 223 BP: 80 A; 34 C; 34 G; 73 T; 2 other:

Query Match 13.6%; Score 172.8; DB 18; Length 223;
 Best Local Similarity 88.7%; Pred. No. 9.2e-33;
 Matches 197; Conservative 0; Mismatches 24; Indels 1; Gaps 1;
 QY 223 GCAGTACATCAATTAATTTGTAAGATTAATTAAGAGACGCTCTATAGACGCA 282
 Db 1 gcaatgacataataattgttaactagaagataataaagaagacgctctatagagacga 60
 QY 283 ATTGAAGGTTTGATTTTATGTTCTGTTAGTAAGATCATCATATGCGCTATAGTA 342
 Db 61 attgaaggtttgattttaaagttctgttagtaagaatcatacatacgaatgcctatagta 120
 QY 343 CTCAGATTATTAATTAATTAACCGTCATTAAATGTTTATTTTGAAGAAACATAGTATCA 402
 Db 121 ctcaagantataataaattaaacgcgtcataaattgttttcgaaaaacattgttaacca 180
 QY 403 TTTTAATGTAAGTGAACATA-CTACGTACTCAATAATCTAT 443
 Db 181 tttnaatgttagtggacctaactacgttctccaataatcat 222
 RESULT 11
 Q10165
 ID Q10165 standard: DNA: 5567 BP.
 XX
 AC Q10165;
 XX
 DT 19-MAR-1991 (first entry)
 DE B.subtilis rib (riboflavin) operon.
 XX
 KM Vitamin B2; ds.
 XX
 OS Bacillus subtilis.
 XX
 FH Key Location/Qualifiers
 FT RBS 352..357
 FT CDS 364..680
 FT /tag= a
 FT /number= ORF 6
 FT 708..748
 FT /tag= c
 FT /standard_name= Rho-independent terminator
 FT 771..799
 FT /tag= d
 FT /label= p1 Promoter
 FT 1034..1067
 FT /tag= e
 FT /standard_name= Rho-independent terminator
 FT 1084..1095
 FT /tag= f
 FT 1101..2186
 FT /tag= g
 FT /number= ORF 5
 FT 1422..1433
 FT /tag= h
 FT 2178..2194
 FT /tag= i
 FT 2197..2844
 FT /tag= j
 FT /number= ORF 4
 FT 2528..2556
 FT /tag= k
 FT /label= p2 Promoter
 FT 2843..2854
 FT /tag= l
 FT 2859..4055
 FT /tag= m
 FT /number= ORF 3
 FT 4071..4085
 FT /tag= n

```

FT CDS 4088..4552
FT /tag= o
FT /label= B-riboflavin synthase gene
FT promoter 4545..4574
FT /tag= p
FT /label= P3 Promoter
FT RBS 4653..4659
FT /tag= q
FT CDS 4665..5039
FT /tag= r
FT /number= ORF 2
FT terminator 5038..5090
FT /tag= s
FT /standard_name= Rho-independent terminator
FT CDS complement (5567..5053)
FT /tag= t
FT /number= ORF 1
XX PA EP405370-A.
XX 02-JAN-1991.
XX 22-JUN-1990; 90EP-0111916.
XX 22-JUN-1989; 89US-0370378.
XX (HOFF ) HOFFMANN-LA ROCHE AG.
XX Perkins JB, Pero JG, Sloma A;
XX WPI; 1991-008756/02.
XX DR P-PSDB; R10569-75, R11221.
XX DNA encoding riboflavin biosynthetic proteins - used to produce
XX recombinant bacteria for prodn. of riboflavin
XX Disclousure; Fig 3; 72pp; English.
XX The sequence or fragments thereof may be expressed in a transformed
XX bacterial expression system for the efficient production of
XX riboflavin (vitamin B2).
XX Sequence 5567 BP; 1751 A; 1096 C; 1297 G; 1423 T; 0 other;
SQ
Query Match 8.1%; Score 103; DB 12; Length 5567;
Best Local Similarity 60.6%; Pred. No. 1.3e-15;
Matches 166; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
OY 729 AATTATGATTAATCGATTCAACTTCCAAATATGATGACAGGTCANACAGGTGTTATC 788
DB 1111 attatagaagctgacctagatcttgcgaagcagggcggaagacagaccgataccatc 1170
OY 789 CACCCGTTGGCGCTGTTGATGTAATGAAGTAGAGATTGTTGATGTCACACTTGA 848
DB 1171 cgctgcgcgcgtctgttcgtaaaagacgaacaaatgtcgaatgagcgccatttaa 1230
OY 849 GAAAAGGTGACACAGCATGGCGAGGTTCAAGCACTTGATATGCGACACANATGCTGAG 908
DB 1231 aatacgtgtaaacctcatgagaagatcattccatcattatgctggaagcactgcagaag 1290
OY 909 GTCGACGATTTATATATTAGTTAGTAGACCATGATGTCATTTGTTCAACACACCCCTGTG 968
DB 1291 gtgccgacatttaacgttaacacgcgaacccgtgcagccattacggaataaacccgcattg 1350
OY 969 TTAACAAAATTAATGATTTGTAAGATAGCANAACT 1002
DB 1351 cagaattgattatcaactctgtatcaataaagaat 1384
RESULT 12
X81947
ID X81947 standard; DNA: 5567 BP.

```

```

XX AC X81947;
XX 10-SEP-1999 (first entry)
XX B. subtilis rib operon nucleotide sequence.
XX DE Riboflavin; open reading frame; ORF; structural gene; promoter;
XX KW vitamin B2; Bacillus subtilis; rib operon; ds.
XX OS Bacillus subtilis.
XX FH Key Location/Qualifiers
XX CDS 1..5567
XX /tag= a
XX /note= "reading frame 1 translated protein (Y21801)"
XX CDS 2..5567
XX /tag= b
XX /note= "reading frame 2 translated protein (Y21802)"
XX CDS 3..5567
XX /tag= c
XX /note= "reading frame 3 translated protein (Y21803)"
XX US925538-A.
XX 20-JUL-1999.
XX 24-AUG-1998; 98US-0138775.
XX 11-SEP-1990; 90US-0581048.
XX 22-JUN-1989; 89US-0370378.
XX 21-APR-1992; 92US-0873572.
XX 06-FEB-1995; 95US-0384626.
XX 24-AUG-1998; 98US-0138775.
XX (HOFF ) ROCHE VITAMINS INC.
XX Erdemberger T, Hatch RT, Hermann T, Perkins JB;
XX PI Pero JG, Sloma A;
XX WPI; 1999-418271/35.
XX DR P-PSDB; Y21801, Y21802, Y21803.
XX Use of bacterial strains for the over production of riboflavin
XX Example 2; Fig 3A-S; 61pp; English.
XX The invention relates to a method of producing riboflavin that comprises,
XX providing a recombinant bacterium containing an exogenously introduced
XX bacterial ORF (open reading frame) 3 (or ORF5) structural gene and an
XX exogenous promoter, culturing the bacterium, and recovery of the product.
XX The method is useful for the production of large quantities of riboflavin
XX (vitamin B2). The present sequence represents the complete nucleotide
XX sequence of B. subtilis rib operon (also see X81948 for the antisense
XX sequence).
XX Sequence 5567 BP; 1751 A; 1096 C; 1297 G; 1423 T; 0 other;
SQ
Query Match 8.1%; Score 103; DB 20; Length 5567;
Best Local Similarity 60.6%; Pred. No. 1.3e-15;
Matches 166; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
OY 729 AATTATGATTAATCGATTCAACTTCCAAATATGATGACAGGTCANACAGGTGTTATC 788
DB 1111 attatagaagctgacctagatcttgcgaagcagggcggaagacagaccgataccatc 1170
OY 789 CACCCGTTGGCGCTGTTGATGTAATGAAGTAGAGATTGTTGATGTCACACTTGA 848
DB 1171 cgctgcgcgcgtctgttcgtaaaagacgaacaaatgtcgaatgagcgccatttaa 1230
OY 849 GAAAAGGTGACACAGCATGGCGAGGTTCAAGCACTTGATATGCGACACANATGCTGAG 908

```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 19, 2000, 05:00:11 : Search time 111.68 Seconds
(without alignments)
1718.530 Million cell updates/sec

Title: US-08-978-456-3

Perfect score: 1269

Sequence: 1 AANCCCAATCCNATGTGGGA.....GTAAACAAGATGCTATAG 1269

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 262060 seqs, 75620727 residues

Total number of hits satisfying chosen parameters: 524120

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/1na/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/1na/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/1na/5C.COMB.seq:*
4: /cgn2_6/ptodata/2/1na/5D.COMB.seq:*
5: /cgn2_6/ptodata/2/1na/5.COMB.seq:*
6: /cgn2_6/ptodata/2/1na/5PCTUS.COMB.seq:*
7: /cgn2_6/ptodata/2/1na/5PCTUS1.COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1250	98.5	1269	5	US-08-978-456-3 Sequence 3, Appl
2	352.2	27.8	1029	5	US-08-978-456-1 Sequence 1, Appl
3	352.2	27.8	3336	5	US-08-977-554-7 Sequence 7, Appl
4	352.2	27.8	3336	5	US-08-978-456-7 Sequence 7, Appl
5	80.4	6.3	3417	4	US-08-978-458-7 Sequence 7, Appl
6	80.4	6.3	3417	4	US-08-978-454-7 Sequence 7, Appl
7	64.4	5.1	1230	4	US-08-741-327E-14 Sequence 14, Appl
8	45.2	3.6	7218	4	US-08-232-463-14 Sequence 14, Appl
9	41.6	3.3	19124	4	US-08-487-826B-13 Sequence 13, Appl
10	38	3.0	731	2	US-08-451-405A-2 Sequence 2, Appl
11	37.6	3.0	1980	4	US-08-804-794-1 Sequence 1, Appl
12	37.2	2.9	1079	4	US-07-781-353-1 Sequence 1, Appl
13	37.2	2.9	19124	4	US-08-487-826B-13 Sequence 13, Appl
14	37	2.9	4248	5	US-08-678-614-1 Sequence 1, Appl
15	37	2.9	4248	5	US-08-678-614-1 Sequence 1, Appl
16	36.4	2.9	2317	5	US-08-465-795-1 Sequence 5, Appl
17	36.2	2.9	2781	5	US-08-749-522-5 Sequence 5, Appl
18	35.8	2.8	8854	5	US-09-053-522-4 Sequence 4, Appl
19	35.6	2.8	8854	5	US-08-724-394A-12 Sequence 12, Appl
20	35.4	2.8	665	4	US-08-883-795A-36 Sequence 36, Appl
21	35.2	2.8	3656	4	US-08-232-463-8 Sequence 8, Appl
22	35.2	2.8	3688	1	US-08-232-463-9 Sequence 9, Appl
23	35.2	2.8	4659	1	US-08-232-463-10 Sequence 10, Appl
24	35.2	2.8	4818	1	US-08-232-463-11 Sequence 11, Appl
25	35.2	2.8	4821	1	US-08-232-463-12 Sequence 12, Appl
26	35.2	2.8	4824	1	US-08-232-463-13 Sequence 13, Appl

27	35.2	2.8	6601	1	US-08-232-463-15 Sequence 15, Appl
28	35.2	2.8	7218	1	US-08-232-463-14 Sequence 14, Appl
29	35.2	2.8	8710	1	US-08-480-882B-3 Sequence 3, Appl
30	35.2	2.8	8710	2	US-08-480-210-3 Sequence 3, Appl
31	35.2	2.8	9019	2	US-08-480-882B-4 Sequence 4, Appl
32	35.2	2.8	9019	2	US-08-480-210-4 Sequence 4, Appl
33	35.2	2.8	9454	1	US-08-232-463-3 Sequence 3, Appl
34	35.2	2.8	9454	1	US-08-232-463-4 Sequence 4, Appl
35	35.2	2.8	9890	1	US-08-232-463-18 Sequence 18, Appl
36	35.2	2.8	9916	1	US-08-232-463-17 Sequence 17, Appl
37	35.2	2.8	9917	1	US-08-232-463-16 Sequence 16, Appl
38	35.2	2.8	10408	1	US-08-232-463-6 Sequence 6, Appl
39	35.2	2.8	10408	1	US-08-232-463-7 Sequence 7, Appl
40	34.8	2.7	665	4	US-08-883-795A-36 Sequence 36, Appl
41	34.8	2.7	51952	5	US-08-947-823-1 Sequence 1, Appl
42	34.6	2.7	1890	7	5312912-3 Patent No. 5312912
43	34.6	2.7	3600	5	US-08-894-731-1 Sequence 1, Appl
44	34.6	2.7	51952	5	US-08-947-823-1 Sequence 1, Appl
45	34.4	2.7	1890	7	5312912-3 Patent No. 5312912

ALIGNMENTS

RESULT 1
US-08-978-456-3
Sequence 3, Application US/08978456
Patent No. 6010881
GENERAL INFORMATION:
APPLICANT: Palmer, Leslie M.
TITLE OF INVENTION: No. 6010881el ribc
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dechert Price & Rhoads
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
CITY: Philadelphia
STATE: PA
COUNTRY: US
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,456
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dickinson, Todd Q
REGISTRATION NUMBER: 28,354
REFERENCE/DOCKET NUMBER: P50444-9
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-994-2252
TELEFAX: 215-994-2222
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1269 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-978-456-3

Query Match 98.5%: Score 1250: DB 5: Length 1269:
Best Local Similarity 100.0%: Pred. No. 0:
Matches 1269: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
QY 1 AANCCCAATCCNATGTGGAGGAGNAATCAATCAATCCCGGANNCCCAATCAAGTTAA 60

```
|||||
Db 1 AANCCAAATCCNATTTGGAGSNAATCCAAATCAATNCCGGANNCCCAATCCAAAGTTAA 60
Oy 61 TTAAGTCCAGGTTTGGAAACATTAACCAAAATATGATTCGGATGAGTCAATGNCANCG 120
Db 61 TTAAGTCCAGGTTTGGAAACATTAACCAAAATATGATTCGGATGAGTCAATGNCANCG 120
Oy 121 GGTGTAATTAATACGAATGNTGTGNAAATGATAGTAANCAAGTTGGCATACAGTAT 180
Db 121 GGTGTAATTAATACGAATGNTGTGNAAATGATAGTAANCAAGTTGGCATACAGTAT 180
Oy 181 CTCNATTAATGATGCTTTAGAAATCGAATGAGCAACAGCGCATGTCATCAATAAT 240
Db 181 CTCNATTAATGATGCTTTAGAAATCGAATGAGCAACAGCGCATGTCATCAATAAT 240
Oy 241 TTGTACTAGAGATTAATAAGAGAAACGCTATAGAGACGAATTTGAAGTTTGAATTTA 300
Db 241 TTGTACTAGAGATTAATAAGAGAAACGCTATAGAGACGAATTTGAAGTTTGAATTTA 300
Oy 301 ATGTCTGTTAGTAAGAAATATATCAATGATGATCCATATGTCATGATTAATTAAT 360
Db 301 ATGTCTGTTAGTAAGAAATATATCAATGATGATCCATATGTCATGATTAATTAAT 360
Oy 361 AAAACCGTCATTAATTTGTTTTTGAAGAAACATATATGATTTTAAATGATTTGACA 420
Db 361 AAAACCGTCATTAATTTGTTTTTGAAGAAACATATATGATTTTAAATGATTTGACA 420
Oy 421 TACTAGTACTCAAAATATCTATTAACAATTTTCAATATATCTTTGCGGCGACGGTGAA 480
Db 421 TACTAGTACTCAAAATATCTATTAACAATTTTCAATATATCTTTGCGGCGACGGTGAA 480
Oy 481 ATTCCCAACCGGACGATTAATAAGCCTGCGACCTGCTAATATGTTCAATATAGTGGCG 540
Db 481 ATTCCCAACCGGACGATTAATAAGCCTGCGACCTGCTAATATGTTCAATATAGTGGCG 540
Oy 541 ATCTAGTGAATTTCTAGACCCGACAGTTAAAGTCTGATGAGGAGAAAGATTTAATAT 600
Db 541 ATCTAGTGAATTTCTAGACCCGACAGTTAAAGTCTGATGAGGAGAAAGATTTAATAT 600
Oy 601 CGACAAAGATTAATGATGCGTATTTGTAATAATGTTGACAAATAGCTTATTTAAGCATAA 660
Db 601 CGACAAAGATTAATGATGCGTATTTGTAATAATGTTGACAAATAGCTTATTTAAGCATAA 660
Oy 661 ATTTTTCCTTGGATCTTAATCAATGATGAGGATTTTGTGTTATAGAGTGATCA 720
Db 661 ATTTTTCCTTGGATCTTAATCAATGATGAGGATTTTGTGTTATAGAGTGATCA 720
Oy 721 TTTGAGTCAATTTATGATTAATGCGATTCAGTTCCAAATATGTTACAAAGTCAANACAG 780
Db 721 TTTGAGTCAATTTATGATTAATGCGATTCAGTTCCAAATATGTTACAAAGTCAANACAG 780
Oy 781 TGTAAATCCACCGCTTGGCGCTGTGTGATTAATGAGGTAGAGATTTGTTATGTCG 840
Db 781 TGTAAATCCACCGCTTGGCGCTGTGTGATTAATGAGGTAGAGATTTGTTATGTCG 840
Oy 841 AACCTTGAGAAAAAGGTGAGAAAGCATGCGGAGTTCAAGACATTGATATGCGCAACANAA 900
Db 841 AACCTTGAGAAAAAGGTGAGAAAGCATGCGGAGTTCAAGACATTGATATGCGCAACANAA 900
Oy 901 TGTGTAAGGTGAGAGATTTATATACGTTAGAGCCATAGTACATTTTGGTTCAACACC 960
Db 901 TGTGTAAGGTGAGAGATTTATATATACGTTAGAGCCATAGTACATTTTGGTTCAACACC 960
Oy 961 ACCCTGTGTTAAACAAATTAATGATTTGAAGATAGCANAAAGTATTAACNCAACANAAG 1020
Db 961 ACCCTGTGTTAAACAAATTAATGATTTGAAGATAGCANAAAGTATTAACNCAACANAAG 1020
Oy 1021 ACAAATTCCTTGAACACATGGGTATAGACAGTTACGGGGCTCCACGGTATTTGAGGG 1080
Db 1021 ACAAATTCCTTGAACACATGGGTATAGACAGTTACGGGGCTCCACGGTATTTGAGGG 1080
Oy 1081 TTGAATTTGGTGAATGAGAAAGGCGATCACAATATACAAAGCTTTTAAAGCA 1140
|||||
```

```
Db 1081 TTGAATTTGGTGAATGAGAAAGGCGATCACAATTAACAAAGCTTTTAAAGCA 1140
Oy 1141 AAAAGCAAGCAACTTGGCCACAAAATTAACAGTGAAGTCTGAAGTTAGATGGGTA 1200
Db 1141 AAAAGCAAGCAACTTGGCCACAAAATTAACAGTGAAGTCTGAAGTTAGATGGGTA 1200
Oy 1201 AACAAAGCAATTTGATATGAGCAAAAGTCAATGATTAACAAAGGTTAAACAGA 1260
Db 1201 AACAAAGCAATTTGATATGAGCAAAAGTCAATGATTAACAAAGGTTAAACAGA 1260
Oy 1261 TGTCTATAG 1269
Db 1261 TGTCTATAG 1269

RESULT 2
US-08-978-456-1
; Sequence 1, Application us/08978456
; Patent No. 6010881
; GENERAL INFORMATION:
; APPLICANT: Palmer, Leslie M.
; TITLE OF INVENTION: No. 6010881el ribc
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; City: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/978,456
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, Todd O
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: P50444-9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2252
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1029 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-978-456-1

Query Match 27.8%; Score 352.2; DB 5; Length 1029;
Best Local Similarity 88.4%; Pred. No. 3,4e-87;
Matches 473; Conservative 0; Mismatches 44; Indels 18; Gaps 8;
```

```
Db 121 GGTGACAAGCATGGGAGGTTCAAGCACTGTATATGACCAACAATAATGCTGAAGGTGCG 180
Qy 914 ACGATTATATAGCTTAGAGCCATGTGATCTTTGGTCAACGCCACCGCTGTAC 973
Db 181 ACGATTATATAGCTTAGAGCCATGTGATCTTTGGTCAACGCCACCGCTGTAC 240
Qy 974 AAAATTATGATTGATAGATAGCAANAAGTAGATTATACNCAANAGCAATTCGCTTAG 1033
Db 241 AAAATTATGATTGATAGATAGCAANAAGTAGA-TACGCACAANAAGCAATT-CGTTAG 298
Qy 1034 ACACACATGGGTGATGAGACGTTACGGGGCTCCACGGTATTGAGGTTGAATGCGTTG 1093
Db 299 ACACACAT-GGTGATGAGACGTTACGGGGCTCCACGGTATTGAGGTTGAATGCGTTG----- 351
Qy 1094 GATGATGACGGGCGCATCTCATTTATACCAAGACTTTTAAAGCAAAAGCAAGCAA 1153
Db 352 GATGATGACGGGCGCATCTCATTTATACCAAGACTTTT---AAAGCAAAAGCAAGCAA 408
Qy 1154 CTGCGCACAAATTAACATGTAAGAGTCTTGAAGAGTTAGATGGTAAACAAGCGAATT 1213
Db 409 CT--GCCACAATTACATGTAAGATC-TGCAAGTTAGATGTAACAAGCG---AAT 462
Qy 1214 GATAATGACAAAGTCAATGATTACTTAACAAGAAGTTAAACAAGATGCTATA 1268
Db 463 GATAATGACAAAGTCAATGATTACTTAACAAGAAGTTAAACAAGATGCTATA 517

RESULT 3
US-08-977-554-7
; Sequence 7, Application US/08977554
; Patent No. 5891672
; GENERAL INFORMATION:
; APPLICANT: Palmer, Leslie M.
; APPLICANT: Fedon, Jason C.
; APPLICANT: Warren, Richard L.
; APPLICANT: Traini, Christopher M.
; APPLICANT: Wang, Min
; APPLICANT: Jaworski, Deborah D.
; APPLICANT: Mooney, Jeffrey
; APPLICANT: Debouck, Christine
; APPLICANT: Zhong, Yiyi
; APPLICANT: Black, Michael
; TITLE OF INVENTION: r1ba
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/977.554
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/02318
; FILING DATE: 19-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, O. Todd
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: P50444-07
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215/994-2252
; TELEFAX: 215/994-2222
;
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
```

```
LENGTH: 3336 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-977-554-7

Query Match 27.8%; Score 352.2; DB 3; Length 3336;
Best Local Similarity 88.4%; Pred. No. 5.8e-87;
Matches 473; Conservative 0; Mismatches 44; Indels 18; Gaps 8;

Qy 734 ATGGATTATGCGATTCACTTCCAAATATGTTAGTACAGGTCCAGAGTGTATATCCACC 793
Db 1 ATGGATTATGCGATTCACTTCCAAATATGTTAGTACAGGTCCAGAGTGTATATCCACC 60
Qy 794 GTTGGCGCTGTTGATGTTAATGAAGTATGATTTGTTGATGTTGTCACACTTGAAGAAA 853
Db 61 GTTGGCGCTGTTGATGTTAATGAAGTATGATTTGTTGATGTTGTCACACTTGAAGAAA 120
Qy 854 GGTGACAAGCATGGGAGGTTCAAGCACTTGTATGGGACACAANAATGCTGAAGGTGCG 913
Db 121 GGTGACAAGCATGGGAGGTTCAAGCACTTGTATTTGGTCACACAANAATGCTGAAGGTGCG 180
Qy 914 ACGATTATATTAAGTATGAGCCATGATCATTTTGGTTCACACCACCCCTGTATTAAC 973
Db 181 ACGATTATATTAAGTATGAGCCATGATCATTTTGGTTCACACCACCCCTGTATTAAC 240
Qy 974 AAAATTATGATTGATAGATAGCAANAAGTAGATTATACNCAANAGCAATTCGCTTAG 1033
Db 241 AAAATTATGATTGATAGATAGCAANAAGTAGA-TACGCACAANAAGCAATT-CGTTAG 298
Qy 1034 ACACACATGGGTGATGAGACGTTACGGGGCTCCACGGTATTGAGGTTGAATGCGTTG 1093
Db 299 ACACACAT-GGTGATGAGACGTTACGGGGCTCCACGGTATTGAGGTTGAATGCGTTG----- 351
Qy 1094 GATGATGACGGGCGCATCTCATTTATACCAAGACTTTTAAAGCAAAAGCAAGCAA 1153
Db 352 GATGATGACGGGCGCATCTCATTTATACCAAGACTTTT---AAAGCAAAAGCAAGCAA 408
Qy 1154 CTGCGCACAAATTAACATGTAAGAGTCTTGAAGAGTTAGATGGTAAACAAGCGAATT 1213
Db 409 CT--GCCACAATTACATGTAAGATC-TGCAAGTTAGATGTAACAAGCG---AAT 462
Qy 1214 GATAATGACAAAGTCAATGATTACTTAACAAGAAGTTAAACAAGATGCTATA 1268
Db 463 GATAATGACAAAGTCAATGATTACTTAACAAGAAGTTAAACAAGATGCTATA 517

RESULT 4
US-08-978-456-7
; Sequence 7, Application US/08978456
; Patent No. 6010881
; GENERAL INFORMATION:
; APPLICANT: Palmer, Leslie M.
; TITLE OF INVENTION: No. 6010881el r1bg
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/978.456
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
```

```
APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
NAME: Dickinson, Todd Q
REGISTRATION NUMBER: 28,354
REFERENCE/DOCKET NUMBER: P50444-9
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-994-2252
TELEFAX: 215-994-2222
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3336 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-978-456-7

Query Match      27.8%  Score 352.2; DB 5; Length 3336;
Best Local Similarity 88.4%  Pred. No. 5.8e-87;
Matches 473; Conservative 0; Mismatches 44; Indels 18; Gaps 8;

QY 734 ATGATTATGCGATTCACCTCCAAATATGCTACAAGGTGTCANACAGGTGTTATCCACC 793
    |||||||
DB 1 ATGATTATGCGATTCACCTCCAAATATGCTACAAGGTGTCANACAGGTGTTATCCACC 60

QY 794 GTTGGCGCTGTGTAGTTAATGAAGTAGAGTATGTTGTTGTTGTCACACTTGACAAA 853
    |||||||
DB 61 GTTGGCGCTGTGTAGTTAATGAAGTAGAGTATGTTGTTGTTGTCACACTTGACAAA 120

QY 854 GTGACAAGATCGGGGCTTCAAGCCTGATATGCGACACANAAATGCTGAAGTGGC 913
    |||||||
DB 121 GTGACAAGATCGGGGCTTCAAGCCTGATATGCGACACANAAATGCTGAAGTGGC 180

QY 914 ACGATTATATACGTTAGAGCCATGATGATTTGGTTCAACACCCTGTTTAC 973
    |||||||
DB 181 ACGATTATATACGTTAGAGCCATGATGATTTGGTTCAACACCCTGTTTAC 240

QY 974 AAATATTATGTTAAGATAGCANAAGTATTTACNCAACANAAGACATTCGTTAG 1033
    |||||||
DB 241 AAATATTATGTTAAGATAGCANAAGTATTTACNCAACANAAGACATTCGTTAG 298

QY 1034 ACACATATGTTGATGAGCTTACGGGGCTCCACGCTATTTGAGGTTGAATTCGTTG 1093
    |||||||
DB 299 ACACATATGTTGATGAGCTTACGGGGCTCCACGCTATTTGAGGTTGAATTCGTTG 351

QY 1094 GATGATGAACGGGCATCACAATTATACCAAAAGACTTTTAAAGCAAAAGCAAA 1153
    |||||||
DB 352 GATGATGAACGGGCATCACAATTATACCAAAAGACTTTTAAAGCAAAAGCAAA 408

QY 1154 CTGGCCACAATAATTCAGTGAAGATNTCTTGAAGTTTAAAGTGGTTAAACAAAGCAATT 1213
    |||||||
DB 409 CT--GGCCACAATAATTCAGTGAAGATNTCTTGAAGTTTAAAGTGGTTAAACAAAGCAAGG--AAT 462

QY 1214 GAATAAGGACAAGAGTCAATGATTAACAAGAGGTAAACAAAGTGTCTATA 1268
    |||||||
DB 463 GAATAAGGACAAGAGTCAATGATTAACAAGAGGTAAACAAAGTGTCTATA 517

RESULT 5
US-08-978-458-7
Sequence 7, Application US/08978458
Patent No. 5932701
GENERAL INFORMATION:
APPLICANT: Palmer, Leslie M.
APPLICANT: Fedon, Jason C.
APPLICANT: Warren, Richard L.
APPLICANT: Kosmatka, Anna L.
APPLICANT: Shilling, Lisa K.
APPLICANT: Stodola, Robert K.
APPLICANT: Knowles, David J. C.
APPLICANT: Black, Michael T.
```

```
APPLICANT: Hodgson, John E.
APPLICANT: Nicholas, Richard O.
TITLE OF INVENTION: r1ba
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESS: Dechert, Price & Rhoads
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103-2793
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,458
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/911,503
FILING DATE: 15-AUG-1997
APPLICATION NUMBER: PCT/US97/14436
FILING DATE: 15-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: Dickinson, O. Todd
REGISTRATION NUMBER: 28,354
REFERENCE/DOCKET NUMBER: P50533-09
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215/994-2252
TELEFAX: 215/994-2222
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3417 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-978-458-7

Query Match      6.3%  Score 80.4; DB 4; Length 3417;
Best Local Similarity 55.2%  Pred. No. 1e-12;
Matches 153; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 729 AATTATGATTTATGCGATTCACCTTCCAAATATGTTACAAAGTGCANACAGGTGTTATTC 788
    |||||||
DB 14 AATTATGATTTATGCGATTCACCTTCCAAATATGTTACAAAGTGCANACAGGTGTTATTC 73

QY 789 CACCCGTGGCGCTGTGTAGTTAATGAAGTAGAGTGTGTTATGTTGTCACACTTGA 848
    |||||||
DB 74 CTATGTTGCGCATTTATTTGTAAGATATTCACATTTTCGACAAAGTTATCATGAT 133

QY 849 GAAGAAGTACAAGCATGCGGAGTTCAAGCACTTGATATGSCACACANAATGCTGAAG 908
    |||||||
DB 134 TTTTGTGGCCACATGCTGAGAGAAATGCTTAAACATGTAAGAAATCCCTGTCG 193

QY 909 GTGCGACGATTTATATACGTTAGAGCCATGTAAGTATTTGGTTCAACACCCTGTG 968
    |||||||
DB 194 GAGCGACGCTTATATGAACACTTGAACCCCTGTTGACCTCGGAAACACACCTCCCTGTA 253

QY 969 TTAACAATAATTTATGTTAGATAGATAGCAANAAGTAGT 1005
    |||||||
DB 254 TAGATGCTATATGATAGTATGATTTACAAAGAGTAGT 290

RESULT 6
US-08-978-454-7
Sequence 7, Application US/08978454
Patent No. 6017728
GENERAL INFORMATION:
APPLICANT: Palmer, Leslie M.
```

```

;
; APPLICANT: Fedon, Jason C.
; APPLICANT: Marten, Richard L.
; APPLICANT: Kosmatka, Anna L.
; APPLICANT: Shilling, Lisa K.
; APPLICANT: Stodola, Robert K.
; APPLICANT: Knowles, David J. C.
; APPLICANT: Black, Michael T.
; APPLICANT: Hodgson, John E.
; APPLICANT: Nicholas, Richard O.
; TITLE OF INVENTION: r1dh
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/978,454
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/911,503
; FILING DATE: 15-AUG-1997
; APPLICATION NUMBER: PCT/US97/14436
; FILING DATE: 15-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, Q. Todd
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: P50533-11
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215/994-2252
; TELEFAX: 215/994-2222
;
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3417 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
; US-08-978-454-7

```

Query Match 6.3%; Score 80.4; DB 5: Length 3417;
Best Local Similarity 55.2%; Pred. No. 1e-12; Matches 153; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

```

QY 729 AATTATGATTAATGCGATCACTCCCAATATGTACAGGTCANACGGTGTATTC 788
DB 14 AATATATGAATTAAGCAATAAACTGCAACAAAAGGGCTGTTACGTCATCCCAATC 73
QY 789 CACCCGTGGCGCTGTGTAGTTAATGAAGTAGAGTTGTGTATGTGTCACACTTGA 848
DB 74 CATATGTTGGCCAAATTAATGTAAGAAATATCATATTCGACAAAGTTATCATGATG 133
QY 849 GAAAGGTACACAAGCATGCGGAGGTTCACAGCACTTATATGCAACAANAATGCTGAAG 908
DB 134 TTTTGTGGTGGCCCATGCTGAGAGAAATGCTTTAAACATGTAGAAAATCCCTGTG 193
QY 909 GTGCGAGATTAATATAGTTAGAGCCATGTATGATTTGTGTTAACAACCCCTGTG 968
DB 194 GAGCGAGCTTATATGTAACACTGAACCTGTGTCTACTTCGGGAAACACCTCCTGTA 253
QY 969 TTAACAAAATTTATGATTTAGATAGCANAAAGTACT 1005
DB 254 TAGATGCTATATGATAGTATGATATTAACAAGTACT 290

```

```

RESULT 7
US-08-741-327E-14
; Sequence 14, Application US/08/741327E
; Patent No. 5925354
; GENERAL INFORMATION:
; APPLICANT: Fuller, et al.
; TITLE OF INVENTION: against Actinobacillus pleuropneumoniae
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: G. Kenneth Smith
; STREET: 300 S. Wacker Drive
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/741,327E
; FILING DATE: October 28, 1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: G. Kenneth Smith
; REGISTRATION NUMBER: 43,135
; REFERENCE/DOCKET NUMBER: 97704-E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-9808
;
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1230 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; US-08-741-327E-14

```

Query Match 5.1%; Score 64.4; DB 4: Length 1230;
Best Local Similarity 51.6%; Pred. No. 1.5e-08; Matches 143; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

```

QY 729 AATTATGATTAATGCGATCACTCCCAATATGTACAGGTCANACGGTGTATTC 788
DB 152 ACTATATGCGCGCGCATTCGCACTGGCAAAACAAGTTAGCTGGACGAATCCCAATC 211
QY 789 CACCCGTGGCGCTGTGTAGTTAATGAAGTAGAGTTGTGTATGTGTCACACTTGA 848
DB 212 CCTGTGTCGGTGTGAATGTCAAAAACGGTGAATCGTCCGAAAGTTCCCATGAAA 271
QY 849 GAAAGGTACACAAGCATGCGGAGGTTCACAGCACTTATATGCAACAANAATGCTGAAG 908
DB 272 AGATTTGTGATGATGCGATGCGGAACGTAATGCCGTTTACATTTGATGAAGAATCTTCCG 331
QY 909 GTGCGAGATTAATATAGTTAGAGCCATGTATGATTTGTGTTAACAACCCCTGTG 968
DB 332 GGGCGAGCTCTTATATGTAACCTTGAAGCCTTGTGTGATCAGCGGCGGACGCGCTGTT 391
QY 969 TTAACAAAATTTATGATTTAGATAGCANAAAGTACT 1005
DB 392 CGATTTATTAATGTAACGAGCATTAATAAAGTAT 428

```

RESULT 8
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.

APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZ9pl-F1s
US-08-232-463-14

Query Match 3.6%; Score 45.2; DB 1; Length 7218;
Best Local Similarity 3.6%; Pred. No. 0.006; Mismatches 164; Indels 0; Gaps 0;
Matches 14; Conservative 211; Mismatches 164; Indels 0; Gaps 0;
QY 869 GAGGTTCAAGCATTGATATGCAACAAAGAGGTTGAGGATTTAATTAG 928
DB 1454 GAGATAGAGATTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1395
QY 929 TTAAGCCATAGTATTGGTTCAACACACCCTGTGTAACAATAATTGATGT 988
DB 1394 RR 1335
QY 989 AAGATAGCANAAGTAGTTTCAACAAAGACATCCGTTGACACATGGGAT 1048
DB 1334 RR 1275
QY 1049 GAGAGTTACGGGCTCCACGTAATTGAGGTTGAATTGGTTGATGAAGGCA 1108
DB 1274 RR 1215
QY 1109 TCACAATTATCAAGAAGCTTTTAAAGCAAAAGCAAGCACTGCCACAAATTA 1168
DB 1214 RR 1155
QY 1169 CAGTGAAGTTCCTGAAGATTGATGGTAAACAAGCAATGATAAGCAAGT 1228
DB 1154 RR 1095
QY 1229 CAATGATTACTAACAAGAGTTAAACA 1257

DB 1094 RR 1066
RESULT 9
US-08-487-826B-13
Sequence 13, Application US/08487826B
Patent No. 5993827
GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knodbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israel, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 19124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-487-826B-13

Query Match 3.3%; Score 41.6; DB 4; Length 19124;
Best Local Similarity 51.0%; Pred. No. 0.09; Mismatches 94; Indels 0; Gaps 0;
Matches 98; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
QY 282 AATGAGGTTGATTTAATGCTGTAGTAAGATCATATCAATGAGTGCCTAGT 341
DB 2072 AAAATTAATGTTATTAATATGACTATAATAATTTTGAATTAACATATGTAAT 2131
QY 342 ACACGATTATTAATTAATAAACCCTCATTAATGTTTTTGAACAATATAGTATC 401
DB 2132 TCATTTAACAGATAAATAAATATATATATATATATATATTAATTAAGTTAT 2191
QY 402 ATTTAATAGTAGTACATAGTACTAGTCAATTAATCTATTAACAATTCATATATAT 461
DB 2192 AATAAATATATATATACATAGATTAATAAATGAAGTTCACTACGTAATATATAT 2251
QY 462 TCTTGGGGCA 473
DB 2252 TATATGTGTCA 2263

ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/781.355
FILING DATE: 19911022
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: UTSH:162
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512-320-7200
TELEFAX: 512-474-7577
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1079 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: unknown
TOPOLOGY: linear
US-07-781-355-1

Query Match 2.9%; Score 37.2; DB 1; Length 1079;

Best Local Similarity 50.0%; Pred. No. 0.39;
Matches 93; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

OY 283 AATTGAAGTTGATTTAATGCTGTTAGTAAAGAAATCATATGAGATGCGCTAAT 341
DB 277 AACAGTAGATTTTATTTATTTTTCATATTTTACTGCTTAATATATGCTTAAAT 218
OY 342 ACTCAGATATTTAATTAACCGTCATTAATGTTTTTTAGAAACATATATGATC 401
DB 217 TAATATTTTATTTCAATTTTTCATATTTTCAATTTTAAATATATATTTTA 158
OY 402 ATTTAATAGTGTGACATGCTACTGCTCAATATCTATACATTTCAATTAAT 461
DB 157 AATCATATATTTTAAATCATATATTTTAAATCATATTTTAAATCATTT 98
OY 462 TCTTTC 467
DB 97 GCCTAC 92

RESULT 13
US-08-487-826B-13/C

Sequence 13, Application US/08487826B
Patent No. 5993827

GENERAL INFORMATION:

APPLICANT: Sim, Kim L.
APPLICANT: Chitais, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487.826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelien, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CPI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 19124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-487-826B-13

Query Match 2.9%; Score 37.2; DB 4; Length 19124;

Best Local Similarity 47.3%; Pred. No. 1.4;
Matches 134; Conservative 0; Mismatches 148; Indels 1; Gaps 1;

OY 123 GTTATTAACACAGAAATGNTGTGNAATGATAGTACAGCAAGTTGCCATACAGTNTCT 182
DB 1520 GTTACTAATATTAATATTTATTTATGCAAAATGTTTGAAGTTTAAAGATTTAT 1461
OY 183 CNATTATATGATGCTTTGAATCGAATGAGCAACACAGCCGCTTACATCAATTAAT 242
DB 1460 TATTTTAAATCAATATATATTTATGAAAAATATATATTTATTTATTTAT 1401
OY 243 GTAACAGATTAATTAAGAGAACGCTCTATAGAGAGATTTAGAGTTGATTTAAT 302
DB 1400 TTTAAAAAAGAGAGACAAATTAATAAAGATTTATATATATATATATTT 1341
OY 303 GTCGTAGTAAATCATATCATGATGCTATAGTACTGATATATTAATTAAT 362
DB 1340 AGATACCGATTTGATTTCTTTTACGCTTGATGTTATACA-AATATATTTATTA 1282
OY 363 AACCGTATTAATGTTTTTTAGAAAAATATAGTATCATTT 405
DB 1281 TAATTTACATATTAATATTTTGTGATACATATCTAATATAT 1239

RESULT 14
US-08-678-614-1

Sequence 14, Application US/08678614
Patent No. 6013507

GENERAL INFORMATION:

APPLICANT: Tomasz, Alexander
APPLICANT: Delencastre, Hermunia
TITLE OF INVENTION: AUXILIARY GENES AND PROTEINS OF
TITLE OF INVENTION: METHICILLIN RESISTANT BACTERIA AND ANTAGONISTS THEREOF
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
STREET: floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk


```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/678,614
FILING DATE: 10-JUL-1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-139
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4248 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Staphylococcus aureus
STRAIN: RUSA 266
FEATURE:
NAME/KEY: ORF
LOCATION: 154..1410
NAME/KEY: ORF
LOCATION: 1497..3500
US-08-678-614-1
```

Query Match 2.9%: Score 37; DB 5; Length 4248;

Best Local Similarity 52.3%; Pred. No. 0.82;

Matches 79; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

```
QY 757 AATATGTACAGTGCACAGTGTATATCCACCCGCTGTGTAGTAAAGA 816
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1943 AACTAAATTACACCTGATTAATGATCAAGCAGCAATCTGATTAATACTAA 2002
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 817 AGTAGAGTTTGTGTATGTGTGACACTTGAGAAAGGTGACAAGCTGGCGAGTTCA 876
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2003 AGCACAAGAAAGTTACAGACGACACAAAGTAAAAAGTTACGACAGATGAAGTTGC 2062
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 877 AGCACTGATATGCGACACANAAATGCTGAA 907
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2063 AGCACTGTACTAAATTAACATGATATAA 2093
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 15
US-08-465-795-1
; Sequence 1, Application US/08465795
; Patent No. 5589355
; GENERAL INFORMATION:
; APPLICANT: Kozumi, Satoshi
; APPLICANT: Yonetani, Yoshiyuki
; APPLICANT: Teshiba, Sadao
; TITLE OF INVENTION: A Process for Producing Riboflavin
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farbow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,795
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,394
FILING DATE: 06-DEC-1993
APPLICATION NUMBER: JP 326578/1992
FILING DATE: 07-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Gaybrick, Robert J.
REGISTRATION NUMBER: 27,890
REFERENCE/DOCKET NUMBER: 04853.0008-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5589 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Corynebacterium ammoniagenes
US-08-465-795-1
```

Query Match 2.9%: Score 37; DB 1; Length 5589;

Best Local Similarity 49.2%; Pred. No. 0.93;

Matches 122; Conservative 0; Mismatches 123; Indels 3; Gaps 1;

```
QY 764 GTACAGGTACANACAGTGTATATCAACCCGCTGCTGTAGTT---AATGAAGT 820
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 672 GTCCGCGGACTACACCCGGAATCCACCGGTGGTCCGTAATTAATTGACTTCGGGT 731
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 821 AGATTTGTGTATGTGTGACACTTGAGAAAGGTGACAAGCTGGCGAGTTCAAGCA 880
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 732 GAGATTTGTGGACACCGGTGACACTACCGCTGGGCGGGGTGCGCAGAGTCCAAAGCT 791
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 881 CTTATATGCGACACANAAATGCTGAGGTGCGACGATTTATATTAGTAGAGCCATGT 940
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 792 CTAGCCGATGCGCGGCGGCAAGCCGAGCGCTACCCCGCTGTGACGCTGGAGCCGCTGC 851
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 941 AGTCAATTTGTTCAACACCAACCCCTGTGTTACAAATTAATTGATTAGATGACANAA 1000
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 852 CGGCATACCGCGCCGACGAGCGCTGCACGACGCTTATTAATGAGCGGCGCATCAAGAT 911
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1001 GTAGTATT 1008
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 912 GTCCTTTT 919
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: November 19, 2000, 05:00:48
Job time: 8885 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 19, 2000, 03:36:00 ; Search time 1232.02 Seconds
(without alignments)
6368.388 Million cell updates/sec

Title: US-08-978-456-3

Perfect score: 1269
Sequence: 1 AANCAACCAATCCNATGTGGA.....GTTAACAAGATGCTATAG 1269

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7189864 seqs, 3091403243 residues

Total number of hits satisfying chosen parameters: 14379728

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

EST: *
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: em_est40:*
42: em_estfun:*
43: em_esthum1:*

44: em_esthum2:*
45: em_esthum3:*
46: em_esthum4:*
47: em_esthum5:*
48: em_esthum6:*
49: em_esthum7:*
50: em_esthum8:*
51: em_esthum9:*
52: em_esthum10:*
53: em_esthum11:*
54: em_esthum12:*
55: em_esthum13:*
56: em_esthum14:*
57: em_esthum15:*
58: em_esthum16:*
59: em_esthum17:*
60: em_esthum18:*
61: em_esthum19:*
62: em_esthum20:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estom:*
68: em_estrov1:*
69: em_estrov2:*
70: em_estp1:*
71: em_estp12:*
72: em_estp13:*
73: em_estp14:*
74: em_estp15:*
75: em_estp1:*
76: em_estro2:*
77: em_estro3:*
78: em_estro4:*
79: em_estro5:*
80: em_estro6:*
81: em_estro7:*
82: em_estro8:*
83: em_estro9:*
84: em_estro10:*
85: em_estro11:*
86: em_estro12:*
87: em_estro13:*
88: gb_gss1:*
89: gb_gss2:*
90: gb_gss3:*
91: gb_gss4:*
92: em_gss1:*
93: em_gss2:*
94: em_gss3:*
95: em_gss4:*
96: gb_gss5:*
97: gb_gss6:*
98: gb_gss7:*
99: gb_gss8:*
100: gb_gss9:*
101: em_gss5:*
102: em_gss6:*
103: em_gss7:*
104: em_gss8:*
105: em_gss9:*
106: em_gss10:*
107: em_gss11:*
108: gb_gss10:*
109: gb_gss11:*
110: em_gss12:*
111: gb_gss12:*
112: gb_gss13:*
113: gb_gss14:*
114: gb_gss15:*
115: gb_gss16:*
116: gb_gss17:*


```

Db 738 TAGRRKRRRTTMRKRRRDRDTRMDADADTRADRRRRRGGDAGAGKRTGRRRRDRBA 679
QY 507 TCGGACCTCTAATATGTTTCAATATTAGTGGCTGATCTAGTAGATTCTAGAG---CCGA 563
Db 678 TMDRDAMWADAMWTTTDTDDMDKRDORRRRGGARRRRRTTARAAMDWTWAMDMW 619
QY 564 CAGTTAACTGAGTGGGGAAGAAGTGAATTATTCGACAAGATTAATGAGCTATT 623
Db 618 DKTRADRRDRADWADWTATRRKADRDMAKARARARDRRAARADRRRTTKKTTTATW 559
QY 624 TGTAAAAATGTGCAAAATAGCCTTATTAACGATAAATTTTCTCTTCATCTTAAT 663
Db 558 TTMAAAAMWAMWAMWATTTATWTTTWTWTTTWTWTTTWTWTTTAAWMAWMTATMA 499
QY 684 TCATGATGTCAGAGATTTTGTTTATAGAGGTCATCTTGAATTTATGAGTATAG 743
Db 498 WTAAAAAAMWMAWMTTWTWTTTWTWMTAAWMTWMTWTTTWTWMAATTTT 439
QY 744 CGATTCACCTTCCAAATATGTCACAGGTCANACAGGTGTTATCCACC 793
Db 438 TWTWTTWMAATTTATTTTWTWTTWMAAAMWMTWMTATATKCCCCC 389

```

```

RESULT 2
LOCUS AVA26745 420 bp mRNA EST 23-MAY-2000
DEFINITION AVA26745 Lotus japonicus young plants (two-week old) Lotus
ACCESSION AVA26745
VERSION AVA26745.1 GI:7785990
KEYWORDS EST
SOURCE Lotus japonicus.
ORGANISM Lotus japonicus
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
Fabales; Fabaceae; Papilionoideae; Lotus.
REFERENCE 1 (bases 1 to 420)
AUTHORS Asanizu,E., Nakamura,Y., Sato,S. and Tabata,S.
TITLE Generation of 7137 non-redundant expressed sequence tags from a
JOURNAL legume, Lotus japonicus
MEDLINE DNA Res. 7 (2), 127-130 (2000)
COMMENT Contact: Yasukazu Nakamura
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: ynakamu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/Plant/.

```

```

FEATURES
SOURCE location/Qualifiers
1..420
/organism="Lotus japonicus"
/db_xref="taxon:34305"
/clone="MM071b02_r"
/clone.lib="Lotus japonicus young plants (two-week old)"
/rev_stage="young plants (two-week old)"
/note="vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI; Isolate-Miyakojima Mg-20"
BASE COUNT 98 a 72 c 132 g 118 t
ORIGIN

```

```

Query Match 5.4%; Score 68.4; DB 18; Length 420;
Best Local Similarity 55.4%; Pred. No. 1.8e-05;
Matches 129; Conservative 0; Mismatches 104; Indels 0; Gaps 0;
QY 773 CANACAGTGTATATGACCCCGTGGCGCTGTGTATGATTAATGAGTGAATGTTGT 832
Db 1 CAGACCAAGCCCAATTCATTTGGGGGTGTGTATTTGTAAGATGCGAAATATGTTGGT 60
QY 833 ATTGTGCACTTGAAGAAAGGTGACAAGCATGCGAGGTCACGACTTGTATGCGA 892
Db 61 CAAGGTGTCACCTTAACAGGAGGAGCCTCATGCTGAGGTGTTGCTTAAGAGATGCG 120

```

```

QY 893 CACANAATGCTGAAGTCGACGATTTATATTACGTTAGACCAGTAGTATTTGGT 952
Db 121 GGGGATTTGCGAGAGCGTGCACGGCTTATGTAGTTGAGCCTCTAATCATTTTGGT 180
QY 953 TCACACACCCCTGTGTTACAAATATTATGATTGTAGATAGCANAGTACT 1005
Db 181 AGACTGACGCTGTCTGACGCTTAATTCAGCCAAAGTGAAGAGTGCT 233

```

```

RESULT 3
LOCUS CNS0155H 1001 bp DNA GSS 26-JUL-1999
DEFINITION Drosophila melanogaster genome survey sequence SP6 end of BAC
BACN13C23 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL105023
VERSION AL105023.1 GI:5617037
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1001)
GENOSCOPE.
Direct Submission
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage;
BP 191 91006 EVRY cedex - FRANCE (E-mail: seqrefgenoscope.cns.fr)
Web : www.genoscope.cns.fr)

```

```

COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billand at CPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.

```

```

FEATURES
SOURCE location/Qualifiers
1..1001
/organism="Drosophila melanogaster"
/plasmid="pBelobAC11"
/db_xref="taxon:7227"
/clone.lib="DrosBAC"
/clone="BACN13C23"
/note="end : SP6"
BASE COUNT 266 a 219 c 134 g 150 t 232 others
ORIGIN

```

```

Query Match 4.7%; Score 60.2; DB 121; Length 1001;
Best Local Similarity 34.4%; Pred. No. 0.0013;
Matches 117; Conservative 61; Mismatches 162; Indels 0; Gaps 0;

```

```

QY 109 AATGNCANCGGTGTTAATACTAGCAATGNTGTGNAATGATAGTAGANCAAGT 168
Db 656 AAANVNMAAARCAARRRRAAAAAAARAGDARAARAAAGRAAARABAAAAATA 715
QY 169 CCGATACAGTNTCTCATTTATGATGCTTTGATCGATGATGACACACAGCGCAGT 228
Db 716 AWWAATTAARWARARBARAAAAAATTAWMTAAWMAWMAWMAWMAAATAAAAADTTWAT 775
QY 229 ACAATCAATTAATTTGTAATAGATTAATAAGACGCTCTATGAGAGCAATGGA 288
Db 776 WMAATTAATAAAMWMAAATAATTTATTAAMMAAATAAATAATTTWMAAAMWTAMA 835
QY 289 GGTGATTTTATGTCGTGTTAGAAATCATATCATAGATGAGATGCTTACTACTACA 348
Db 836 AAAAAAATTAATTTTAAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMA 895
QY 349 TTAATTAATTAATTAATTAATTTTATTTTATTTTATTTTATTTTATTTATTTA 408
Db 896 AATTAATAAAMWMTWMAWMAWMAWMTWTTTWTWMAWMAAATAATTAATTAATTA 955

```

Y	409	ATGAGTTCACATCTACCTACTCAATATCTATCTA	448
		: : : : : : :	
Db	956	WATATAAATATAAAAAATWMAAAWATATWTTAAAAA	995
RESULT	4		
CNS00E07			
LOCUS			
DEFINITION	CNS00E07 1101 bp DNA GSS 04-JUN-1999		
ACCESSION	Drosophila melanogaster genome survey sequence TET3 end of BAC:		
VERSION	BACR29P01 of RPc1-98 library from Drosophila melanogaster (fruit		
KEYWORDS	fly), genomic survey sequence.		
SOURCE	AL069440		
ORGANISM	AL069440.1 GI:4949583		
AUTHORS	GSS.		
TITLE	fruit fly.		
JOURNAL	Drosophila melanogaster		
COMMENT	Eumetazoa; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 1101) Genoscope. Direct Submission Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : sequef@genoscope.cns.fr - Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org/TheBDGP/Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammosser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPc1-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2: cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm . location/Qualifiers		
FEATURES	1..1101		
source	/organism="Drosophila melanogaster"		
	/db_xref="taxon:7227"		
	/clone_lib="RPc1-98"		
	/clone="BACR29P01"		
	/note="end : TET3"		
BASE COUNT	366 a 66 c 104 g 351 t 214 others		
ORIGIN			
Query Match	4.7% ; Score 59.8 ; DB 121 ; Length 1101 ;		
Best Local Similarity	35.1% ; Pred. No. 0.0016 ;		
Matches	202; Conservative 77; Mismatches 296; Indels 1; Gaps 1.		
Y	190	ATGATGCTTTAGATCGAATGAGCAACAACAGCCAGGTACATCAATATATTGTGACTA	249
		: :	
Db	442	AKAATATTTTGGAAAGATATAAAAAARATAAWTAATTAATTAATTAATAA	501
Y	250	GAGCATATAAAGACAACGCTCTATAGAGACCAATGAAGCTTGATTTATATGCTGTT	309
		:	
Db	502	WTTTAAAAAAWMAAATAAATAAATCAAMTWTAAATTAATTTAAATAAAMWATKTT	561
Y	310	AGTAGAATCATATGATGAGATGCTATAGTACACAGATTTATTAATTAACCGTC	369
		: :	
Db	562	TTTATWMAATATAAATAAATTTATTTTATTTATTTATTTAAWMAAAWMAAAWMT	621
Y	370	ATTAATTTGTTTTTAGAAAACATATAGTATCATTTTAAATAGTACTGCATACGTGA	429
		: :	
Db	622	ATAAHTWTTAAATTTATATWTTAAWMTTTTAAATATATTTAAWTTAAWTTTAAWA	681
Y	430	CTCAAAATATCTATACAAATTTTCATATATATCTTTTCGCGGACAGGTCAAATTC	489
		: : : : : : :	

Db	682	TTTTTAAMTATTTAAAAAAMWATTWAAAAAAMAAAVTTTWAATAAATAAATAAATAATTTT	741
Qy	490	CGGCACTAATAATAAGCCTCGGACCGCTGAATATGTTCATATTAAGTGACTATCAGCA	549
Db	742	TTTTTAAATAAMWTTTAAAATTAAWTATWATAAANAAMWTTWMAAAAAATNTWATAAMATTW	801
Qy	550	GATCTTAGAGCCGACAGTAAAGTCTGGATGGAGGAAAGAATGTAATATTCAGCAAGA	609
Db	802	-WKATTAATWATAAMWTTTAAATAATWARTAAATWATAATWTTTTTTTAAWTTTAAWA	860
Qy	610	TAACTAGCGATTTGTAAAAATGTGTCAAAATAGCCTTATTTAAACAATAATTTTCTC	669
Db	861	TTTTTTTMMTWAMAAMWTTAAATWAAATTAAWTTAAAAAAMATGMTTWTAAATWMTGTA	920
Qy	670	CTTGCATCTTAATCATGATGTGGAGATTTTTTTTGTATAGAGTGATCATTTAGCA	729
Db	921	TTTTTTTAAATTTAAATTCATATTAATAAAMWTTTATGTATTAAMAAATRTMWTGTATTAA	980
Qy	730	ATTATGATTTATGCCATTCACCTTCCAATATGCT	765
Db	981	AATATWGRATATTKAAAAAAAAAAATAATWATAAAMART	1016
RESULT	5	CNS00390	1101 bp DNA GSS 03-JUN-1999
LOCUS		Drosophila melanogaster genome survey sequence TERT3 end of BAC #	
DEFINITION		BACR08M12 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.	
ACCESSION		AL063931	GI:4941788
VERSION		AL063931.1	GI:4941788
KEYWORDS		GSS.	
SOURCE		fruit fly.	
ORGANISM		Drosophila melanogaster	
		Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;	
		Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;	
		Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
REFERENCE		Genoscope.	
AUTHORS		Direct Submission	
TITLE		Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : fr	
JOURNAL		BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)	
COMMENT		- Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammeter in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization along with the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.	
FEATURES		location/Qualifiers	
SOURCE		1..1101	
		/organism="Drosophila melanogaster"	
		/db_xref="taxon:7227"	
		/clone_id="RPCI-98"	
		/clone="BACR08M12"	
		/note="end : TERT3"	
BASE COUNT		260 a 155 c 97 g 385 t 204 others	
ORIGIN			
Query Match		4.6% ; Score 59 ; DB 121 ; Length 1101 ;	
Best Local Similarity		37.6% ; Pred. No. 0.0024 ;	
Matches 120 ; Conservative		49 ; Mismatches 150 ; Indels 0 ; Gaps 0 ;	

Query Match	4.6%	Score 57.8	DB 113	Length 337
Best Local Similarity	54.6%	Pred. No. 0.0042		

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	french-Constant, R.H., Waterfield, N., Burland, V., Perna, N.T., Daborn, P.J., Bowen, D., and Blattner, F.R.	A genomic sample sequence of the enteropathogenic bacterium <i>Photobacterium luminescens</i> W4: potential implications for virulence Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)	Department of Biology and Biochemistry University of Bath South Building, Bath BA2 7AY, UK Tel: (44) 1225 826621 Fax: (44) 1225 826779 Email: bssr@icbath.ac.uk	This is one of 2,122 random reads from the M13 library. For annotation of identified clones (BLASTX, BLASTN and mapping to E. coli K12 genome) please see french-Constant et al. 2000, Nucleic Acids Res.
2	Seq primer: M13 Forward			
3	Class: Shotgun.			
4	Location/Qualifiers			
5	1. 337			
6	/organism="Photobacterium luminescens"			
7	/strain="W4"			
8	/db_xref="taxon:29488"			
9	/clone="PLG02461"			
10	/clone_lib="Photobacterium luminescens strain W4 M13 library"			
11	/dev_stage="primary phase variant"			
12	/note="genomic DNA from strain W4 was size selected (1-2 kb) and then cloned into M13 Janus."			
13	86 a	60 c	93 g	97 t
14	ORIGIN			1 others

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	ORGANISM	SOURCE
AO991504	337 bp DNA	GSS	08-AUG-2000			
LOCUS	RfC02461 Photobacterium luminescens strain W4 M13 library					
DEFINITION	Photobacterium luminescens genomic clone PLG02461, DNA sequence.					
ACCESSION	AO991504					
VERSION	AO991504.1	GI:9650098				
KEYWORDS	GSS.					
ORGANISM	Photobacterium luminescens.					
SOURCE	Photobacterium luminescens					
ORGANISM	Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae; Photobacterium.					
REFERENCE	1 (bases 1 to 337)					
AUTHORS	french-Constant, R.H., Waterfield, N., Burland, V., Perna, N.T., Daborn, P.J., Bowen, D., and Blattner, F.R.					
TITLE	A genomic sample sequence of the enteropathogenic bacterium Photobacterium luminescens W4: potential implications for virulence Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)					
JOURNAL	Contact: french-Constant RH					
COMMENT	Department of Biology and Biochemistry University of Bath South Building, Bath BA2 7AY, UK Tel: (44) 1225 826621 Fax: (44) 1225 826779 Email: bssr@icbath.ac.uk					
	This is one of 2,122 random reads from the M13 library. For annotation of identified clones (BLASTX, BLASTN and mapping to E. coli K12 genome) please see french-Constant et al. 2000, Nucleic Acids Res.					
	Seq primer: M13 Forward					
	Class: Shotgun.					
	Location/Qualifiers					
	1. 337					
	/organism="Photobacterium luminescens"					
	/strain="W4"					
	/db_xref="taxon:29488"					
	/clone="PLG02461"					
	/clone_lib="Photobacterium luminescens strain W4 M13 library"					
	/dev_stage="primary phase variant"					
	/note="genomic DNA from strain W4 was size selected (1-2 kb) and then cloned into M13 Janus."					
	86 a	60 c	93 g	97 t	1 others	
	ORIGIN					

Matches	113:	Conservative	0:	Mismatches	94:	Indels	0:	Gaps	0:
QY	729	AATTATGATTAATGCGATTTCACTTCCAAATATGTTCAAGTCANACAGTGTATATC							788
Db	125								184
QY	125	AATATATGCGCGCTCTTGGAAATTAGCGCTCAAGGGCGCTTACACGTCACCTATATC							184
Db	789	CACCGTGGCGCTGGTTAGTATGAAGTAGGATGTTGGTATGTCACACTGA							848
QY	185								244
Db	185	CCACGTTAGTGTGTGTATGTTGCGTGTATGGCAGATTAATTGGGAGAGTTTCAATTAC							244
QY	849	GAAGAAGTACAGACATCGAGGCTTCAAGCAGCTGATATGCAACAMACAMATGCTGAAG							908
Db	245								304
QY	909	GTCGACGATTTATATTACGTTAGAC							935
Db	305	GCGCACTGTATTATGTCACGCTTGACC							331
RESULT	7								
LOCUS	CNS00LT2	1101 bp	DNA						GSS
DEFINITION	Drosophila melanogaster genome survey sequence TERT end of BAC: BACR48P19 of RPC1-98 library from Drosophila melanogaster (fruit fly) , genomic survey sequence.								14-JUN-1999
ACCESSION	AL078714	GI:5102004							
VERSION	AL078714.1								
KEYWORDS	GSS.								
SOURCE	fruit fly.								
ORGANISM	Drosophila melanogaster								
REFERENCE	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.								
AUTHORS	1 (bases 1 to 1101)								
TITLE	Genoscope.								
JOURNAL	Direct Submission								
COMMENT	Submitted (11-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)								
FEATURES	Source								
FEATURES	1..1101								
FEATURES	/organism="Drosophila melanogaster"								
FEATURES	/db_xref="taxon:7227"								
FEATURES	/clone_lib="RPC1-98"								
FEATURES	/clone="BACR48P19"								
FEATURES	/note="end : TERT"								
BASE COUNT	469 a 6 c 69 g 151 t 406 others								
ORIGIN									
Query Match	4.68: Score 57.8; DB 121: Length 1101;								
Best Local Similarity	24.48: Pwd. No. 0.0044:								
Matches 158: Conservative 188: Mismatches 256: Indels 6: Gaps 2:									
QY	200	AGAATCGAATAGACACACAGCGAGTTACATCAATATATTTGTAAGTACAGAGATAATA							259
Db	400	AAA							459

[illegible]

DB	661	NTNNTAAVNTNNNNNNNTTNNVNTNVNNVNTNTNANNNNNNNNNNNNNNNNNNNTTTNNNT	622
OY	669	CCCTTGCACTCTTAATTCATGATGAGCATTTTTCGTTTAAGAGCTATCATTTGAC	726
Db	621	TNNTNNNNNNNNVVTTTTTNNANNNTNNNVVNTTNNNNVTNNVNVAGVNNNNNNNNNV	564
RESULT 12			
LOCUS	CNS00396	1101 bp	GSS
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR08K10 of Rpci-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.		
ACCESSION	AL063921		
VERSION	AL063921.1	GI:4941778	
KEYWORDS	GSS.		
SOURCE	fruit fly.		
ORGANISM	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
REFERENCE	1 (bases 1 to 1101)		
AUTHORS	Genoscope.		
TITLE	Direct Submission		
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)		
COMMENT	- Web : www.genoscope.cns.fr Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammeter in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named Rpci-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2: cn bw sp. the same strain used for the BDGP's pl and Est libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .		
FEATURES	Location/Qualifiers		
source	1..1101		
	/organism="Drosophila melanogaster"		
	/db_xref="taxon:7227		
	/clone_1db="Rpci-98"		
	/clone="BACR08K10"		
	/note="end : TET3"		
BASE COUNT	201 a	64 c	131 g
ORIGIN	202 t	503 others	
Query Match	4.3%:	Score 54.6:	DB 121; Length 1101;
Best Local Similarity	19.1%:	Pred. No. 0.023:	
Matches	105;	Conservative	207; Mismatches 237; Indels 0; Gaps 0;
OY	221	GCGCAGTACATCATATATTGTAACAGAGAATATAAGAGAACGCTATAGACAC	280
Db	392	GCGMATATAMMMWMTTTTTTTTMAAAMAAATATTTMAAAMAAATTTWMAAA	451
OY	261	CAATTGAAGCTTGATTTAATGTCGTAGTAGAATCATCATAGATGCTTAG	340
Db	452	WAAAAMATTAWTTTATWMAAAMAAATTTTWTWTTTATWTTATWTTTMMW	511
OY	341	TACTCAGTTATTAATTAATAACCGCATTAATGTTTGTTTAGAGAACATATAGAT	400
Db	512	TTAAAMAAAAAAMAAAAAMAAAMATATAATWTTTWTWTTTAAATTAAMCMAAM	571
OY	401	CATTTAATAGTAGTGCATCTAGTCTAGTCAATAATATATACAACTTCATATATA	460
Db	572	YHTTTTYTHYHTTWTGTYTMTNHYHNTYHNAAHNTTWYTWYTAAMHMTNHWMTMAWH	631

[illegible]

RESULT	13
CNS0182P	
LOCUS	
DEFINITION	CNS0182P 1101 bp DNA GSS 26-JUL-1999
	Drosophila melanogaster genome survey sequence SP6 end of BAC
	BACN37D10 of DrosBAC library from Drosophila melanogaster (fruit
	fly), genomic survey sequence.

ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE
AL108811	AL108811.1	GI:5629115	fruit fly, GSS.	<i>Drosophila melanogaster</i>	
				Euarthropoda; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroideae; Drosophilidae; Drosophila.	
				(bases 1 to 1101)	

REFERENCE AUTHORS TITLE JOURNAL	COMMENT
1 (bases 1 to 1101) Genoscope. Submitted Submission Directed (22-jul-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr)	- Web : www.genoscope.cns.fr - Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk - This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billard at CEPH (Centre d'etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBeloBAC11.

FEATURES	SOURCE	LOCATION/QUALIFIERS
		1. 1101
		/organism="Drosophila melanogaster"
		/plasmid="pbeloBAC11"
		/db_xref="taxon:7127"
		/clone_11b="DrosBAC"
		/clone="BACN37D10"
		/note="end : Sp6"
BASE COUNT	274 a	268 c 128 g 73 t 358 others
ORIGIN		

Query Match	4.3%	Score	54.6	DB	121	Length	1101
Best Local Similarity	21.7%	Pred	No. 0.023				
Matches	126	Conservative	209	Mismatches	243	Indels	2
						Gaps	2

Oy	137	AAATGTTGTGNNAAATGATAGTACGACAGTGGCCATACAGTTCCTCCNATTATATATATGC	196
Db	533	AAAAAAAAAAAAAAAAAAGGCGACARACAGCGRGRGVGAGRGATATATAAAAAAAAAAADD	582

Oy	197	TTTGATTCGATTAGCAACAAACGGCAGTTCATCATCATTAATTTGTACTAAGAAGT -	255
Dd	583	TKDWDMDRGAAAARAANAADAAARMAAKWTAAAAAABAAAAAAAAAAAAAAAAAGTTKTTA	642
Oy	256	AATAAAGAGAACCGCTCTATAGAGACGAATGAAGGTTTTATTTATGTCTGTAGTAG	315
Dd	643	AAAAAAGGCMKTCKGKAALDDGCAAAADDMWRGRGRRAARAARAANAADKKRAAAAA	702
Oy	316	AATCATATCAATGAGATGCCTATAGTACTCAGATTATTAATTAAATTAACAACGTCATTAT	375
Dd	703	AAAAAAAAAAAAADAGCRKKMKKGDKDGKATTAAMAGRKMDGATATMTMTDAMDYT	762
Oy	376	TGTTTTTTAGAAAAACATATAGTATCATTTTAATTAATTAAGTTACACTACTACTCAA	435
Dd	763	WKAATTDIDKRAAAGRKRKPARRKTABDGGRRATRRRAAAGSKRRAPACARRARAARRA	822
Oy	436	TAACTCTAACAAATTTTCATATATATCTTCGCGGAGCGGTGAATCCCAACGCCGAG	495
Dd	823	DDRDMDMAAAAAAAAAAAAAATTTWRDRKMDMDMDYT RMDDDTTAAWMDAARARRRRR	882
Oy	496	TAAATTAAGCCTGGACCCTGCTAATATGTTTCATTTAGTGCCTGATCTAGTAGATTCT	555
Dd	883	RRRRRRARARAARAADDYDTKD RMADATTTXDTTWTTDDDDMDKAKRD- RWMAAAD	941
Oy	556	AGAGCCGACAGTTAAAGCTGCGATGGGAGACAAGATGTTAAATTTCGACAACAAGATAAGT	615
Dd	942	GAMWKPRBARDDMAATAKDDDKDWKMGWGSGRGKGDKKRMWDKGTGKKDDDDMDKWTNR	1001
Oy	616	AGCGTATTTGTAANAATAGTACAAATAGCGTTATTTAACGATAAATTTTCTCCTTGC	675
Dd	1002	DMMMWMTRTKMDWMDWDGGRGRGWTRRMCWAMWRAAMAARDYDTGKDYRTPAOKDRKFTDT	1061
Oy	676	ATCTTAATTCATGATGGAGAGATTTTGTGTTTATACAGG	715
Dd	1062	TKRDGDDMRKDRKDKRRDKGDDKTKKADTAHYDDDAARPM	1101

RESULT	14
CNS008H1	
LOCUS	
DEFINITION	961 bp DNA GSS 03-JUN-1999
ACCESSION	CNS008H1
VERSION	
	Drosophila melanogaster genome survey sequence TENG end of BAC #
	BA0R17001 of RPCI-98 library from Drosophila melanogaster (fruit
	fly), genomic survey sequence.
	AL051882
	AL051882.1 GI:4933734

KEYWORDS	SOURCE	ORGANISM	REFERENCE
fruit fly.	CSS.	Drosophila melanogaster	1 (bases 1 to 961)
		Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta.	
		Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;	
		Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	

COMMENT	JOURNAL
Submitted (02-JUN-1999) Genoscope - Centre National de Sequençage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segreit@genoscope.cns.fr - Web : www.genoscope.cns.fr)	Determination of this BAC-end sequence was carried out as part of a

melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazuhiro Osoegawa and Aaron Mammeter in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

Location/Qualifiers

BASE COUNT	ORIGIN	246 a	85 c	73 g	458 t	99 others
Query Match	Best Local Similarity	4.3%	Score 54	DB 121	Length 961	
Matches 174	Conservative	43	Mismatches	243	Indels	0
					Gaps	0
1	961					
/organism="Drosophila melanogaster"						
/db_xref="taxon:7227"						
/clone_11b="RPC1-98"						
/clone="BACR17001"						
/note="end : TET3"						
282	AATTGACGTTGATTTAATGTCGTGTAGTAAAGAAATCATATCATGAGATGCTATGCTACT	341				
103	AAATTTTAAATTTAAATTTAATTTTATTTTGTGTCMTCTWTATTTAATTTAAATTTTATTTT	162				
342	ACTGAGATTATTTAATTAATTAACCGTCTTATTTGTTTGTAGAAACATATAGTATC	401				
163	TTTTTTTTTTTTTTAACTTACATACATATATATTTCTTTAAATAGAAATTTACTTTTATA	222				
402	ATTTTAAATGTAGTGCATCTAGTACATCTGACATATCATATCATATATCATATATAT	461				
223	AATTTTATTTTAAATTTTAAATTTTATTTATTTTATTTTATTTATTTATTTATTTATTT	282				
462	TCTTTCGGGGCAGGCTGAATTCCTCAACCGCAGTAAATTAAGCCTCGACCTGCTAATA	521				
283	TTTTTACTTATTTATKATATATATATATTTTTCATCTWTATTTATTTTATTTTATTTTAA	342				
522	TGTTTCATTTATGTCGCGATCTAGTACATCTGACGCGACGATTAAGTCTGATG	581				
343	TTTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT	402				
582	GAGAAGACATGTTAATTTATTCGACAAAGATATGATGCGTATTTGTTAAATGTGTACAA	641				
403	ATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	462				
642	TAGCTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT	701				
463	TTTTTTTATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT	522				
702	TGTTTATAGAGTATCATTTGATGATCATTTATGATTA	741				
523	TTATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT	562				
RESULT 15						
CNS008X3						
LOCUS						
DEFINITION						
ACCESSION						
KEYWORDS						
ORGANISM						
SOURCE						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
COMMENT						

Arnon Mammosset, Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of *Drosophila* DNA provided by the BSGP from the isogenic strain y2. cn bw^{sp}, the same strain used for the BSGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

```
FEATURES
source      location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_11b="RPI-98"
/clone="BACR18L14"
/notes="end : 17"
251 a      151 c      169 g      239 t      291 others
BASE COUNT
BRIGIN
```

Query Match	4.2%;	Score 53.8;	DB 121;	Length 1101;
Best Local Similarity	33.2%;	Pred. NO. 0.035;		
Matches 130;	Conservative 80;	Mismatches 178;	Indels 4;	Gaps 1.

QY	73	TTTGGAAACATTTCCCAATATGTTCCGATGAGGTCATGCAATGCANCGGTGTTATAAC	132
Db	652	TNTHWAAAWTTTWTATATATTTBMBABRBRABTSNMAAAASSAAABNSVBDBAATATAAA	711
QY	133	TACGAAATGNTGTGMAATGATAGTAGANCAAGTGGCATACAGTNTCTCATTTATG	192
Db	712	AABVATATBBBAAAAABABBSBVSASASAAANMSANMBBSBVNAAAVAAAAATATA	771
QY	193	ATGCTTAGAATGCATGAGCATGAGCAACACACCGCAGTTACATCATATTTGTAACATAGA	252
Db	772	AAARARAKKAAAAAARAAAAARAKRRDDAARAAATAAAAAGAAADAAADAAAAAKAKAA	831
QY	253	GATPATAAAGAGAACCGCTGATAGACA---GGAATGAGGTTGATTTAAATCTGCT	308
Db	832	TAKKAAKAAATATAKKGKRRKRKAACKTTTAAKAAWMAAAATTTATWTTTMAATWMAAD	891
QY	309	TAGTAAGAATCATATCAATGAGATGCCCTAGTACTACAGATTATATTAATTAATAAACGT	368
Db	892	AAAAAAWADAAAAAAAKADPAATATATAAKAKAAATATAMWADATAAADTAAAWATAA	951
QY	369	CATTAAATGTTTTTTTAGAAAACATATAGTATCATTTTAAATGTAGTGCATCTACTAGCT	428
Db	952	TATKWTATATWTTAAAAAWMAAKARTTATATMAAWTTTATATAMGATGAAWMAWMAAAGT	1011
QY	429	ACTCAATATATCTATACAACTTTCATATATA	460
Db	1012	AAAAAAKAKARTAAWMTTATATWATGTMWMAWMAA	1043

Search completed: November 19, 2000, 03:36:09
Job time: 12218 sec